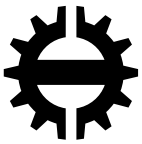


TAMPERE UNIVERSITY OF TECHNOLOGY

TARJA ÄIJÄNEN

INVESTIGATION OF INTERACTIONS BETWEEN CHOLESTERYL ESTER TRANSFER PROTEIN AND ANACETRAPIB BY USING MOLECULAR SIMULATIONS

Master of Science Thesis

Examiners: Ilpo Vattulainen, Hannu Eskola
Examiners and topic approved in the
Faculty of Computing and Electrical
Engineering council meeting on
9 November 2011

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Electrical Engineering
ÄIJÄNEN, TARJA : Investigation of interactions between cholesteryl ester transfer protein and anacetrapib by using molecular simulations
Master of Science Thesis, 79 pages, 18 Appendix pages
February 2012
Major: Medical Physics
Examiners: Professor Ilpo Vattulainen, Professor Hannu Eskola
Keywords: anacetrapib, atherosclerosis, CETP, cholesterol levels, computer simulations, helix X, inhibitors

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality in Western societies. The most important cause behind these diseases is a complex condition affecting arterial blood vessels denoted as atherosclerosis. Atherosclerosis is inflicted and restrained by lipoproteins circulating in blood. More specifically, high levels of low density lipoprotein (LDL) have been found to correlate positively and high levels of high density lipoprotein (HDL) inversely with the risk of atherosclerosis.

High LDL levels can be reduced by the use of statins, a strategy that has successively reduced the rate of CVD. However, a significant residual risk still remains since atherosclerosis is a multifactorial disease. Because low HDL levels are an independent risk factor, elevating HDL has become one the most promising strategies in the fight against CVD. Clinical trials have shown that this can be achieved through the inhibition of cholesteryl ester transfer protein (CETP), which transports neutral lipids between different lipoprotein fractions. A novel molecular agent, anacetrapib, has been found to meet these requirements but the precise inhibitory mechanism remains to be elucidated.

In this study the interactions between CETP and anacetrapib were examined using atomistic molecular dynamics simulations. Extensive structural and functional analysis were performed for both particles. The obtained results point towards the important regulatory roles of helix X and phospholipids during the lipid exchange process. These structures were found to experience considerable conformational fluctuations induced by the drug, indicating the possible capability of anacetrapib to inhibit the functions of CETP.

The performed simulations are pioneering and pave the way for further studies, with an objective to extend the scope of computational studies to gain a much deeper understanding concerning the inhibition of CETP. The novel insight could be used in the development of new molecular agents capable of preventing the progression of CVD.

TIIVISTELMÄ

TAMPEREEN TEKNILLINEN YLIOPISTO
Sähkötekniikan koulutusohjelma
ÄIJÄNEN, TARJA: Kolesteroliesterien siirtäjäproteiinin ja anacetrapib-lääkeaineen välisten vuorovaikutusten tutkiminen molekyylisimulaatioilla
Diplomityö, 79 sivua, 18 liitesivua
Helmikuu 2012
Pääaine: Lääketieteellinen fysiikka
Tarkastajat: Professori Ilpo Vattulainen, Professori Hannu Eskola
Avainsanat: anacetrapib, ateroskleroosi, CETP, helix X, inhibiittorit, kolesterolitasot, tietokonesimulaatiot

Sydän- ja verisuonisairaudet ovat yleisin kuolinsyy teollisuusmaissa. Tärkein tekijä näiden sairauksien taustalla on valtimonkovettumatauti, ateroskleroosi. Ateroskleroosi on verenkiertoelimistön toimintaan liittyvä sairaus, jonka ominaispiirteitä ovat kolesterolin asteittainen kasautuminen sekä paikallisten ateroomaplakkien muodostuminen verisuonten seinämiin. Ateroskleroosin synnyssä merkittävässä roolissa ovat veren lipoproteiinit. Kohonneen matalatiheyksisen lipoproteiinin (low density lipoprotein, LDL) pitoisuuden uskotaan edistävän sairauden syntyä, kun taas kohonneet korkeatiheyksisen lipoproteiinin (high density lipoprotein, HDL) pitoisuudet alentavat sairastumisriskiä.

Kohonnutta LDL-pitoisuutta voidaan alentaa statiineilla, jotka ovat yleisimmin määrättyjä lääkkeitä ateroskleroosin hoidossa. Kuitenkin vain joka kolmannen potilaan kolesteroli alenee niiden avulla tavoitetasoon johtuen ateroskleroosin monitekijäisyydestä. Koska matala HDL-pitoisuus on luokiteltu itsenäiseksi riskitekijäksi, on HDL-tason nostaminen yksi lupaavimmista tavoista sydän- ja verisuonisairauksien ehkäisemisessä. Kliiniset kokeet ovat osoittaneet, että tämä voidaan saavuttaa estämällä kolesteroliesterien siirtäjäproteiinin (CETP) toimintaa. CETP:n toiminnan estämiseen kehitetyn lääkkeen, anacetrapibin, on osoitettu vastaavan vaatimuksia. Lääkkeen tarkka estämismekanismi on kuitenkin vielä selvittämättä.

Tämä tutkimus keskittyy CETP:n ja anacetrapibin välisten vuorovaikutusten tutkimiseen molekyylidynamiikkasimulaatioiden avulla. Molempien molekyylien rakennetta sekä toimintaa tutkitaan laaja-alaisesti. Tutkimustulokset korostavat helix X:n sekä fosfolipidien tärkeää säätelevää roolia lipidien vaihdon aikana. Näiden rakenteiden huomattiin uudelleenjärjestäytyvän anacetrapibin vaikutuksesta, mikä viittaa lääkkeen kykyyn estää lipidien vaihtoa lipoproteiinien välillä.

Tehdyt simulaatiot ovat laajuudessaan ensimmäisiä ja siksi ne tarjoavatkin erinomaiset lähtökohdat tuleville CETP:n toiminnan estämiseen liittyville tutkimuksille. Näitä tietoja voidaan käyttää uusien, sydän- ja verisuonisairauksien etenemistä ehkäisevien lääkkeiden kehityksessä.

PREFACE

This Master of Science thesis has been carried out in the Biological Physics research group at the Institute of Physics at Tampere University of Technology during summer and autumn 2011.

First of all I would like to thank my examiner professor Ilpo Vattulainen for giving me the opportunity to work with a very interesting subject in an encouranging working environment. I wish to thank my other examiner, professor Hannu Eskola, for his advices and suggestions related to this work.

Special thanks I would like to express to my supervisor Artturi Koivuniemi for all his help, ideas and advices throughout the project. I also want to thank all of my other colleagues part of the research group for assisting in the completion of this work. Especially I would like to mention my officemates Dominika, Sanna and Sami who offered help every time it was needed as well as for the many pleasant conversations during the past months.

In the end, I want to thank my family and friends for all their company and support I have been able to enjoy during the work.

Tampere, February 2012

Tarja Äijänen

CONTENTS

SYMBOLS AND ABBREVIATIONS

α_{HB}	Angle between the acceptor-donor-hydrogen in hydrogen bond formation
b_{ij}	Bond rigidity for bond stretching
$\mathbf{b}$	Box vectors for Parrinello-Rahman pressure coupling
β	Determines the relative weight of the direct and reciprocal sums in Ewald summation
β_{ij}	Isothermal compressibility
C^{α}	Alpha carbon, central carbon in amino acids
C_{ij}	Covariance matrix
δ_{ij}	Kronecker delta
e	Unit electron charge
E	Binding energy
ϵ_0	Permittivity in vacuum
ϵ_{ij}, σ_{ij}	Lennard-Jones parameters for particles i and j
ϵ_r	Relative permittivity
f	Dimensional factor of binding energy in secondary structure analysis
$\mathbf{F}_i$	Force acting on particle i
$g(r)$	Spatial distribution function
$g(r_{ij})$	Radial distribution function
$g(\omega)$	Radial angular distribution function
k_B	Boltzmann constant
k_BT	Thermal energy
k_{ij}^b	Force constant for bond stretching
k_{ijk}^{θ}	Force constant for bond angle
k_{ϕ}	Force constant for dihedral angle
λ	Eigenvalue
m_i	Mass of particle i
$\mathbf{m}$	Reciprocal-space vector
M	Sum of particle masses or diagonal matrix
n	Selectable constant for dihedral angle potential
n_x, n_y, n_z	Box vectors
N	Number of particles
μ_{ij}	Scaling matrix for Berendsen pressure coupling
$\mathbf{p}(t)$	Principal component

P_{0ij}	Target pressure
$\mathbf{P}$	Pressure tensor
$\mathbf{P}_0$	Reference pressure tensor
ϕ	Angle between two vectors
ϕ_{ijkl}	Dihedral angle
ϕ_s	Equilibrium angle for dihedral
q_i	Charge of particle i
Q	Parameter for describing the strength of Nosé-Hoover coupling
$\mathbf{r}_{cm}$	Distance from the center of mass of a molecule
r_{HB}	Distance of hydrogen-acceptor in hydrogen bond formation
r_{ij}	Distance between particles i and j
$\mathbf{r}_{ij}$	Vector between particles i and j
$\mathbf{r}_{ij,\mathbf{n}}$	The real distance between charges i and j
R	Transformation matrix
R_g	Radius of gyration
$\rho(r,\theta)$	Number of molecules in the volume element
t	Time coordinate
Δt	Short timestep
T	Temperature
T_0	Reference temperature
τ	Berendsen temperature coupling time constant
τ_p	Berendsen pressure coupling time constant
θ^0_{ijk}	Equilibrium angle between the three particles i, j and k
θ_{ijk}	Angle between the three particles i, j and k
$\mathbf{v}_i$	Velocity of particle i
V	Potential or volume
V_a	Bond angle potential
V_b	Bond stretching potential
V_C	Coulombic potential
V_d	Potential for dihedral angle
V_{dir}	Direct sum term in Ewald summation
V_{LJ}	Lennard-Jones potential between particles i and j
$V(\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_N)$	Complete force field
V_{rec}	Reciprocal sum term in Ewald summation
V_0	Constant term in Ewald summation
$\mathbf{W}$	Parameter determining the strength of the Parrinello-Rahman pressure coupling
x_i	Position of atom i
ξ	Friction term

ABCA1	ATP-binding cassette A1
ABCG	ATP-binding cassette G
apoA-I	Apolipoprotein A-I
apoA-II	Apolipoprotein A-II
apoB-100	Apoliporotein B-100
ASCVD	Atherosclerotic cardiovascular disease
ATP	Adenosine triphosphate
CE	Cholesteryl ester
CETP	Cholesteryl ester transfer protein
CRP	C reactive protein
CVD	Cardiovascular disease
DOPC	Dioleoylphosphatidylcholine
DSSP	Define secondary structure of protein
HDL	High density lipoprotein
HDL-C	High density lipoprotein-cholesterol
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
IC_{50}	Half maximal inhibitory concentration
IDL	Intermediate density lipoprotein
LCAT	Lecithin-cholesterol acyltransferase
LDL	Low density lipoprotein
LDL-C	Low density lipoprotein-cholesterol
LJ	Lennard-Jones
MD	Molecular dynamics
PBC	Periodic boundary conditions
PCA	Principal component analysis
PME	Particle mesh Ewald
PL	Phospholipid
PLTP	Phospholipid transfer protein
RCT	Reverse cholesterol transport
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
SDF	Spatial distribution function
SR-B1	Scavenger receptor class B member 1
TG	Triglyceride
UC	Unesterified cholesterol
VLDL	Very low-density lipoprotein
VMD	Visual molecular dynamics

1. INTRODUCTION

Cardiovascular disease (CVD) is the leading cause of death in Western societies covering all abnormal functions of the heart and blood vessels [1]. These diseases cause incapacity for work, and in addition they also correspond to most of the costs of the public health service. One of the risk factors behind these diseases is the elevated level of low density lipoprotein cholesterol (LDL-C) in circulation. Low density lipoprotein (LDL) particles transport cholesterol and its esters from the liver to peripheral tissues where they are used as the building material of cell membranes and for the production of hormones, vitamin D and bile acids [2]. However, the excess transportation of LDL-C is followed by a complex series of events affecting arterial blood vessels, leading to a condition called atherosclerosis, which is the underlying and the most important cause behind different CVDs [1]. Hence LDL particles are often denoted as "bad cholesterol". The early stage of an atherosclerotic plaque forms when LDLs accumulated in the arterial intima become oxidized and altered in other ways. As a consequence, an inflammatory response is stimulated that fuels the progression of atherosclerosis, eventually leading to formation of atherosclerotic plaques. This gives rise to the narrowing of arteries, rupture of the plaque, clotting, and finally a potential death.

Another plasma lipoprotein denoted as the high density lipoprotein (HDL) possesses important antiatherogenic functionalities. It transports cholesterol from the opposite direction compared with LDL, that is, from peripheral tissues into the liver to be excreted and recycled [2]. Due to this, HDL is denoted as "good cholesterol". Up to date, the most sensitive indicator of CVDs is the relation of LDL-C to HDL-cholesterol (HDL-C). If it is not in accordance with general guidelines (< 3.5 mmol/dL), it eventually leads to the accumulation of LDL particles into arterial wall [3]. In the biggest risk are those who have elevated LDL-C level, decreased HDL-C level, an unhealthy way of life and obesity or genetic genotype for CVDs.

Since the formation of an atherosclerotic plaque takes decades to develop into a phase where vessel narrowing causes clinical complications, better ways of detecting the disease are needed. Several clinical trials have shown that cholesterol-lowering medication effectively reduces the risk of atherosclerosis since the lipoproteins have a significant role in the formation of the plaques [1, 4]. These novel molecular agents have been developed in order to provide everybody with equal possibilities to lower

the levels of cholesterol and, more specifically, to lower the level of LDL-C and to increase the level of HDL-C. These levels can be influenced also with healthy ways of life such as reducing alcohol consumption, cutting down on smoking and following a versatile physical exercise and diet.

In the past decade, lowering the level of LDL-C has been the major target in the prevention of cardiovascular events. This approach has proven to be beneficial and effective for both the primary and secondary prevention of CVDs since numerous useful drugs, such as statins, have been developed. However, a large portion of cardiovascular events cannot be prevented by LDL-C lowering strategies since atherosclerosis is a multifactorial disease. Because low levels of HDL have been identified as an independent risk factor, elevating HDL levels has become one of the most promising strategies in the fight against CVDs. Hence the search for finding even better treatment methods is more and more focusing on HDL.

Several clinical trials have shown that the inhibition of cholesteryl ester transfer protein (CETP) increases the level of HDL-C in the human bloodstream. CETP is a plasma protein that transports cholesteryl esters (CEs) and triglycerides (TGs) between HDL and other lipoprotein particles. As a consequence, HDL is catabolised more rapidly and cholesterol is transported mainly via LDL [5]. Therefore a promising strategy for reducing the risk of CVDs is to develop novel molecular agents that inhibit the lipid transfer functions of this protein. This is not a straightforward issue concerning drug development since the first developed drug, torcetrapib, increased the HDL-C level as was hoped but it also increased the blood pressure as well as the mortality rate [6]. Due to this, all the clinical trials concerning torcetrapib had to be terminated. The development of the drug continued and a real breakthrough was published at the end of the year 2010 when a new variant of the drug, anacetrapib, was found to inhibit CETP without severe cardiovascular events [7].

Despite these promising achievements, the precise mechanism of how this drug inhibits the functions of CETP remains unknown. In addition, the relationship between the inhibition and increases in HDL-C levels is yet to be proven. Therefore the ultimate purpose of this work is to study the interactions between CETP and anacetrapib and to gain more detailed information in this regard. The aim is to tie the observed findings with the inhibitory mechanism of the drug. The interactions are studied using atomistic molecular dynamics (MD) simulations which offer a possibility to examine small scale particles with detailed knowledge on an atomic scale. The performed simulations are pioneering and serve as a solid foundation for future studies concerning the interactions of the whole HDL-CETP-anacetrapib complex. The obtained results can be used for the development of new molecular agents in the fight against the generation and progression of CVDs.

This work is divided into six chapters. The next chapter discusses the biological background relevant to this work, describing the connection between atherosclerosis, levels of cholesterol, inhibition of CETP and anacetrapib. The molecular structures of CETP and anacetrapib are presented in depth. The third chapter explains the classical molecular dynamics method in detail, with an emphasis on the force field and the ensembles used to describe the physical conditions. The fourth chapter contains the initial configurations of the simulated systems as well as the parameters applied in the simulations. In addition, it lists the methods used to analyze the obtained data. The chosen methods are commonly used to analyze the changes in the structure and functions of proteins. The fifth chapter presents the results obtained from the simulations and discusses their physical importance. Finally, the conclusions summarize the results and present ways of how this research can be used in future studies.

2. BIOLOGICAL BACKGROUND

In this chapter, the biological background relevant to this thesis is discussed. First are reviewed the formation and progression of atherosclerosis as a complex form of CVD as well as the role of plasma lipoproteins in this process. Next the structure and role of CETP in the transportation of CEs and TGs between different lipoprotein fractions are discussed. Inhibiting the lipid transfer functions of this protein increases the level of HDL-C and decreases the level of LDL-C in human blood plasma, thus correlating with a reduced risk of CVDs. Inhibition can be obtained through the use of novel molecular agents, and hence anacetrapib is presented as one possible CETP inhibitor.

2.1 Relevance of cardiovascular diseases

CVDs encompass all abnormal functions of the heart and blood vessels and are among the leading causes of death and disease both in Finland and in other western countries [1]. In addition to causing incapacity to work, CVDs are responsible for a great deal of the costs of the public health service. The underlying and the most important cause behind these diseases is atherosclerosis, a multifactorial disease, which narrows the arteries as a cause of cholesterol build up and formation of an atherosclerotic plaque. This prevents the flow of blood to deliver oxygen to tissues and may finally lead to plaque rupture, clotting and potential death. Atherosclerosis develops in the course of several years and has usually proceeded far before the first symptoms appear. Therefore it is important to pay attention to the development of different treatment methods in order to reduce the risk of CVDs.

2.1.1 Cardiovascular diseases in general

The group of CVDs can be divided into the ischaemic heart diseases, cerebrovascular diseases and diseases of the arteries as presented in table 2.1 [1]. The most general symptoms of CVDs include the myocardial infarction (heart attack), heart failure, cerebral infarction (stroke) and cardiac rhythm problems [8]. The probability of these cardiovascular events to occur increases with multiple risk factors since they reinforce each other in their influence on morbidity and mortality. This leads to the reduction of the quality of life and life expectancy.

Table 2.1: Cardiovascular diseases. Modified from [1].

Ischaemic heart diseases	Cerebrovascular diseases	Diseases of arteries
Angina pectoris	Cerebral infarction	Atherosclerosis
Acute myocardial infarction		Aortic aneurysm
Subsequent myocardial infarction		Arterial embolism
Chronic ischaemic heart disease		Arterial thrombosis

The diagnosis, care and follow-up of several risk factors play a central role in the prevention and care of CVDs. The risk factors include different environmental, biochemical and genetic risk factors, see table 2.2 [8, 9]. These factors can be divided into preventable and unpreventable risk factors. The factors that can be prevented or treated include alcohol consumption, smoking, physical inactivity, diet, obesity, diabetes, high blood pressure and levels of cholesterol. The unpreventable risks include the appearance of CVDs in the family, gender and age. The heritability of CVDs has been shown in many family and twin studies [4].

Table 2.2: Risk factors for cardiovascular disease. [1, 4]

Risk factor	Environmental	Physiological or biochemical	Genetic
Alcohol consumption	x		x
Smoking	x		x
Lack of physical activity	x		x
Diet	x		x
Obesity, body composition		x	x
Diabetes and blood glucose levels		x	x
Hypertension		x	x
Low HDL levels		x	x
High LDL levels		x	x
High triglyceride levels		x	x
Thrombotic factors		x	x
Family history of CVD		x	x
Male gender			x
Age			x

The most important risk factor is the elevated level of cholesterol in circulation which is a consequence of the amount and function of the plasma lipoproteins [4]. It has been shown that atypical levels of lipoproteins are a prerequisite for most forms of CVDs. In the group of the biggest risk are those who have elevated levels of LDL-C as well as reduced levels of HDL-C. The general guidelines for these levels are < 3 mmol/L and > 1 mmol/L, respectively, as well as < 5 mmol/L for the total cholesterol level [3]. If the cholesterol levels are consistent with these

recommendations, it helps to keep the heart healthy.

Healthy lifestyle including the cutting down on excessive alcohol consumption and smoking, as well as versatile physical exercise and diet are essential parts of the basic care of CVDs. With the help of sufficient physical exercise the risk of coronary artery disease can be prevented and, in addition, it also has positive effects concerning the reduction of obesity and the diseases associated with it such as the high blood pressure [8, 10]. Healthy and versatile diet lowers the levels of cholesterol when the intake of saturated fat as well as the consumption of sugary and processed foods are avoided [10, 11]. In addition, the diet does not cause costs comparable with medication. Another approach to lower the cholesterol levels is to use pharmacological agents that reduce the production of cholesterol in the liver or reduce the absorption of cholesterol when the removal of cholesterol from tissues increases [12]. Furthermore, the elevated cholesterol content of blood can be reduced with the help of agents which prevent the absorption of bile acids (e.g. resins and guarkumi) or with agents that prevent the formation of fats (e.g nicotinic acid and its derivatives) [12].

The clinical data from multiple trials have demonstrated that in the primary and secondary prevention of cardiovascular events the 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, statins, are the most effective agents to reduce the level of LDL-C [11, 13]. The effect of statins is based on the inhibition of the HMG-CoA reductase when the synthesis of cholesterol in the liver is prevented [14]. As a consequence, the number of LDL-receptors of liver cells increases and LDL-C is transported from the circulation into the liver [15]. Thus the levels of total cholesterol and LDL-C decrease by 25 % to 55 % depending on the agent and dose. Statins are well tolerated medicines. They have been reported to have so-called pleitropic effects, which include the reduction of oxidative stress and vascular inflammation as well as the improvement of the stability of atherosclerotic plaque [16, 17]. The clinical importance of pleitropic effects is still unclear but they have inspired researchers to look for new applications for statins.

Despite their life saving aspect, a significant residual risk of cardiovascular events is still apparent under statin therapy. This is because atherosclerosis is a multifactorial disease and the current medication is not enough to prevent the progression of CVDs. Since low levels of HDL have been identified as an independent risk factor, also the therapy concerning the elevation of HDL-C levels has to be improved. Clinical trials concerning this issue have involved the use of niacin, cholesterylamine and fibrates [10, 11, 18]. Niacin has been shown to increase the HDL-C levels by 15-35 % through inhibiting the secretion of very low density lipoprotein (VLDL) particles, increasing the lipoprotein lipase activity and decreasing the TG levels [19]. Fibrates (e.g. gemfibrozil) have been shown to raise the levels of HDL-C by 10-15 % through the hepatic expression of the main HDL apolipoproteins A-I and A-II (apoA-I and

apoA-II) [19, 20]. These results are promising, yet there still is a strong need to develop even more efficient therapies. One strategy to meet these requirements is to use CETP inhibitors which are a novel way to lower the levels of HDL-C, see Section 2.3 for further details.

2.1.2 Atherosclerosis

CVDs such as the heart disease and stroke usually result from a complex series of events affecting arterial blood vessels leading eventually to a condition called atherosclerosis [1]. This disease is usually denoted as atherosclerotic cardiovascular disease (ASCVD) whose severity can be predicted from the risk factors listed at table 2.2 [21]. Popular descriptions concerning ASCVD suggest that cholesterol gradually builds up on the surface of passive vessel walls [2]. The formed plaque eventually closes off the affected vessel if it grows large enough, thus preventing blood from flowing. If tissue sustains lack of oxygen, it dies causing a stroke or heart attack. As recently as two decades ago, scientists have proposed that cholesterol accumulates within the vessel walls where its components become oxidized and altered in other ways. The altered components stimulate an inflammatory response that fuels the development of ASCVD [2].

The structure of the normal healthy artery is presented in the first cross section in Figure 2.1. The artery consists of three morphologically distinct layers denoted as the intima, the media and the adventitia. The normal intima is a thin region of the vessel wall closest to the bloodstream, consisting of a connective tissue matrix of proteoglycans and collagen [4]. It is the essential layer concerning the generation of atherosclerosis since the early stage of an atherosclerotic plaque forms when excess LDL-C starts to accumulate within the intima, see the cross section two. In addition to its detrimental effects, cholesterol is needed as the building material of cell membranes and to the production of hormones, vitamin D and bile acids [22]. After accumulation, the lipids as well as the proteins composing LDL particles undergo oxidation [23]. The altered components lead endothelial- and smooth muscle cells to secrete chemicals which attract inflammatory cells to attach to the vessel wall and to multiply and mature into active macrophages. Macrophages, also denoted as foam cells, take up the extensively modified lipid particles with the help of various receptors and ligands, see the cross section three [2].

HDL particles are also carried from the circulation into the intima. They are smaller in size than LDLs and do not bind to the extracellular matrix composed of proteoglycans. When HDLs meet foam cells, cholesterol that resides in them starts to flow into HDLs [24]. This is followed by the HDL-mediated cholesterol transport back to the liver thus defining the antiatherogenic role of these particles. If modified LDLs are formed in the intima more than HDLs are able to remove, cholesterol

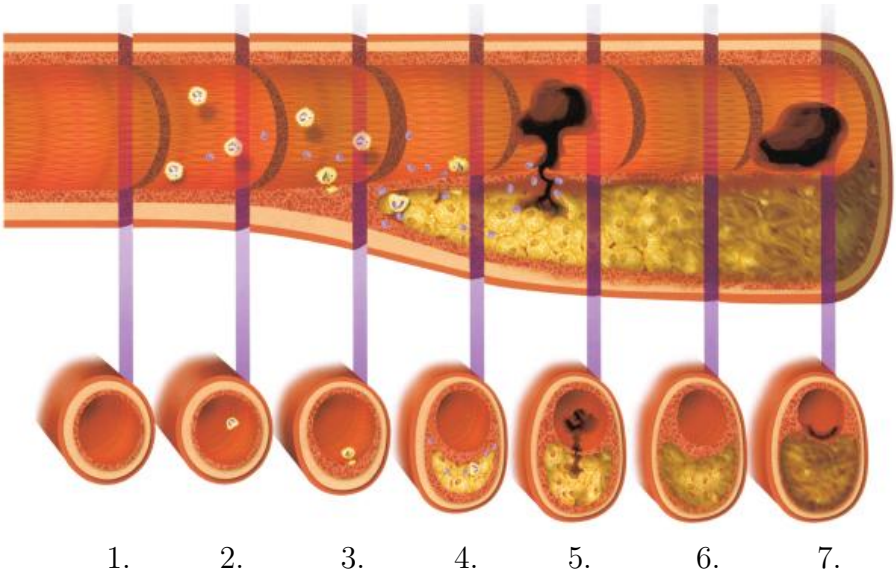


Figure 2.1: The formation of an atherosclerotic plaque. The longitudinal section of the artery (top) depicts the timeline of the initiation and the progression of atherosclerosis from normal artery (1) to atherosclerotic plaque (5, 6, 7). The cross sections of the artery (bottom) describe the evolution of the plaque. See text for further details. Modified from [14].

starts to accumulate also outside the cells [24]. As a result, the chemicals secreted by the endothelial cells attract white blood cells into the intima. The white blood cells together with the foamy macrophages produce many inflammatory mediators that amplify the inflammatory activities in the vessel walls and, as a consequence, atherosclerotic plaque starts to develop [2, 13].

The inflammatory activities change the character of the vessel wall, generating a bigger and more complicated plaque with a fibrous cap over the lipid core, see the cross section four in Figure 2.1. The cap increases the size of the plaque but additionally separates it out from the blood. The integrity of the cap depends on collagen fibers made by the smooth muscle cells which can be degraded by the enzymes secreted by macrophages [13]. If the weakened plaque ruptures as presented in the cross section five, blood encounters the lipid core with proteins able to make the blood to coagulate causing a clot [2]. If the clot is big enough, it will prevent the flow of blood to the heart causing a heart attack, thrombus or stroke. The cross sections six and seven present a healing response caused by the products associated with thrombosis. The healing process leads to an increased collagen accumulation and a growth of smooth muscle cells which can produce the symptoms of stable chest pain [14].

Recently, several clinical studies have suggested a different approach for the formation of the plaque. These studies have reported that the formation of the plaque is more related to the pile up of calcium carbonate on the vessel walls. This

calcification is followed by an inlammatory response fuelled by the accumulation of excess cholesterol. It has been noticed that vitamin D has a central role in this process as well as the fact that there is a link between vitamin D deficiency and the different risk factors of CVDs [25, 26]. Several mechanisms explain the reasons behind this link. First, low levels of vitamin D increase the levels of the parathyroid hormone which in turn increases the risk of inflammation [27]. Second, the supplementation of vitamin D may down-regulate the serum markers of inflammation [26, 27]. However, more clinical studies are needed to confirm these issues but undoubtedly the formation of atherosclerosis is a very complex process.

Despite the controversial suggestions, the plaque formation takes decades to develop into a phase where vessel narrowing causes clinical complications. Due to this, better ways of detecting ASCVD are needed since narrowed coronary arteries are not able to supply enough blood and oxygen to the heart. In addition to healthy lifestyle, several clinical trials have shown that lipid lowering medication reduces the risk of ASCVDs by reducing the cholesterol production of liver or by lowering the absorption of cholesterol due to the significant role of lipoproteins in the formation of plaques.

2.1.3 Lipoproteins

Lipoproteins, the water-soluble colloidal particles transporting insoluble lipids within the human bloodstream [28], are among the most studied structures in biology. The most important reason behind this interest is the crucial role of lipoproteins in the emergence and treatment of different CVDs. However, the understanding about the precise molecular structures and functions of lipoproteins in the nanometer length scales have been a challenging task, but the development of MD simulation methods and computing power has offered added value for the research of lipoproteins over the nanometer length scales as well as over the microsecond time scales.

Lipoproteins are lipid-protein complexes that can be divided into different fractions based on their density. These fractions are denoted from the most sparse to the densest as chylomicrons, VLDL, intermediate density lipoprotein (IDL), LDL and HDL, as presented in table 2.3 [24]. Each particle contains a different amount of phospholipids (PL), unesterified cholesterol (UC), TGs, CEs and specific apolipoproteins bound onto the surface of lipoproteins. From these the PLs, UC and apolipoproteins form an amphiphilic monolayer that surrounds a hydrophobic core consisting of TGs and CEs [24]. The apolipoproteins originate in the liver and intestines and regulate the dynamic processes of lipoprotein assembly by binding to specific enzymes or transport proteins on the cell membranes, thus directing the lipoprotein to the proper site of metabolism [28]. These transfer proteins are called CETP and phospholipid transfer protein (PLTP), which modify the composition of lipoproteins through

neutral lipid transfer [24]. CETP transports TGs, CEs and PLs between lipoproteins whereas PLTP transports only PLs [24]. In addition to size, density and lipid-protein composition, the lipoprotein particles differ from each other also in regard to their individual functions in metabolism.

Table 2.3: Lipoproteins and their composition. The size, the density, the major apolipoproteins and the percentage values of PLs, TGs and CEs are presented [10, 19, 24, 28, 29].

	Chylomicrons	VLDL	IDL	LDL	HDL
Size (nm)	100-500	30-90	25-50	18-28	5-15
Density (g/mL)	< 0.93	0.93-1.006	1.006 -1.019	1.019 -1.063	1.063 -1.210
Apolipoproteins	B-48, C-II, E	B-100, E, C-II	B-100, E	B-100	A-I, A-II
PLs (%)	8	20	22	29	46
TGs (%)	85	50	31	13	5
CEs (%)	3	18	29	50	40

Except for LDL and IDL, all lipoprotein particles are synthesized in the liver. The chylomicrons and VLDL particles transport mainly TGs from the liver to peripheral tissues for utilization [28]. The primary function of HDL and LDL particles is the transportation of cholesterol and its esters in the bloodstream. More specifically, LDL particles transport cholesterol from the liver to peripheral tissues whereas HDL particles transport it from the opposite direction from tissues into the liver to be recycled or excreted [24]. IDL is similar to LDL and it transports both TGs and cholesterol [28].

Low density lipoprotein

LDL is a spherical particle presented in Figure 2.2. The hydrophilic monolayer consists of PL components called phosphatidylcholine and sphingomyelin and a single copy of apolipoprotein B-100 (apoB-100) drawn as dark blue, light blue and grey, respectively [30]. The apoB-100 has an important role together with receptors and different enzymes to maintain the structural integrity and to control the interactions of LDL particles [30]. The hydrophobic lipid core is comprised of TGs and CEs drawn as green and yellow. Both the monolayer and the core contain UC drawn as red.

Although the composition and overall structure of LDL particles are well known, the experiments done to obtain information on the interactions between the surface and core components have led to contradictory results [30, 31]. Due to this, the model of LDL comprising the surface monolayer and the lipid core have not been

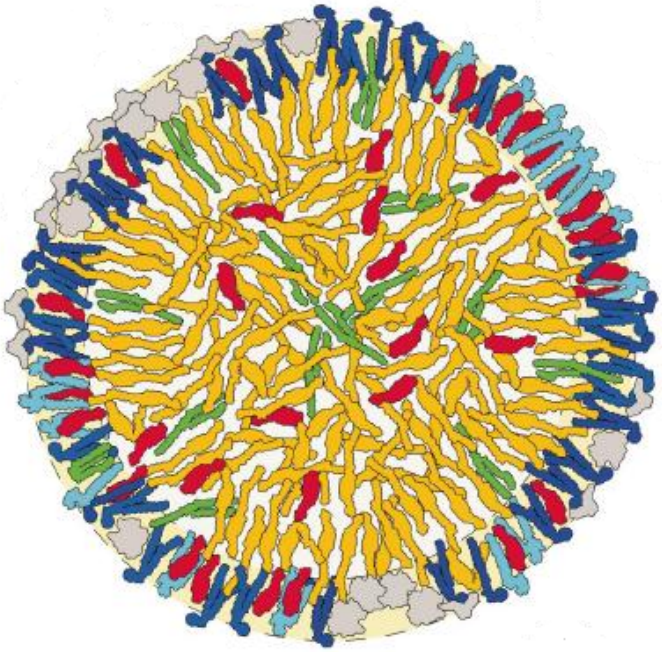


Figure 2.2: Schematic representation of the LDL particle. The molecular components are drawn with different colours: phosphatidylcholines as dark blue, sphingomyelins as light blue, TGs as green, CEs as yellow, UC as red and apoB-100 as grey. Modified from [30].

accurate enough, and therefore a new model has been introduced. Hevonoja et al. [30] described LDL with a three-layer model where the LDL particle has been divided into three radially dissimilar structural layers named as the outer surface layer, the interfacial layer and the core [30]. The outer surface layer consists of PL headgroups whereas the interfacial layer consists of interdigitating core and surface lipids [30]. The pure core consists of molecules that are not in the direct contact with the surface monolayer. The apoB-100 have parts in each of these three layers [30]. However, Murtola et al. [31] found that the distribution as well as the ordering of the lipids in the LDL particle are rather complex, thus suggesting that the three-layer model is inadequate to describe the properties of lipids in LDL particles [31]. Hence additional studies are needed to understand the structural as well as dynamical properties of LDL particles.

LDL as well as IDL particles are not synthesized in the liver as the other lipoprotein particles, but instead the formation of these particles occur during the VLDL metabolism. In this chain reaction VLDLs are first hydrolyzed into IDLs which are then converted into LDL [24]. When the VLDL mediated TG particles enter the capillaries of peripheral tissues they come into contact with lipoprotein lipase, hepatic lipase and endothelial lipase. These components together with an activating protein, apolipoprotein CIII on the VLDL surface, hydrolyze TGs and release fatty

acids for tissue utilization [19, 32]. Due to this, the TG content of VLDL is reduced thus inducing this particle to convert into IDL. Part of IDLs are removed by the liver and part are converted into LDL during a process where a hepatic lipase further degrades the TGs inside IDL [19, 32].

LDL particles are the key players in cholesterol transport and metabolism because they are the main carriers of cholesterol in circulation. They transport cholesterol from the liver and intestines to other tissues which use it to repair the membranes or to produce steroids [2, 33]. In addition, LDLs transport vitamins and hormones to various cell types [33]. Despite these fundamental functions, LDL possesses also detrimental effects. One of these is the excessive transportation and accumulation of cholesterol in the intima which promotes the generation of ASCVDs. Due to this, LDL is often called "bad cholesterol". The relationship between LDL-C and ASCVDs has been established through several animal models and clinical investigations. High LDL levels can be reduced by the use of statins, a strategy that has successively reduced the rate of cardiovascular events. Despite the use of these agents, a significant residual risk remains since nearly two-thirds of cardiovascular events continue to occur [18, 20]. Therefore the focus has shifted beyond LDL-C lowering to include therapies that increase the levels of HDL-C. Indeed, several clinical trials have authenticated the proposal that increased HDL-C levels correlate with a reduced risk of ASCVDs [34, 35].

High density lipoprotein

High density lipoprotein is a CE-rich particle characterized by the smallest size and the highest density of the plasma lipoproteins, see table 2.3. The structure presented in Figure 2.3 is similar with the structure of LDL despite that HDLs a constitute remarkably heterogeneous group differing in density, size, lipid composition and apolipoprotein content. Due to this, HDL particles can be fractionated into five subclasses ranging from small, dense HDL_{3c}, HDL_{3b} and HDL_{3a} to larger HDL_{2a}, HDL_{2b} and discoidal pre-β HDL [36]. The main apolipoproteins of HDL are apoA-I and apoA-II having amphiphilic alphahelical structures that help them to attach to the surface of the lipid matrix [37]. These apolipoproteins are required for the efficient cholesterol efflux from macrophages as well as to regulate the transportation of lipids into and out of cells [38]. In addition, HDL contains several other proteins and enzymes of which lecithin-cholesterol acyltransferase (LCAT) has an important role in cholesterol esterification [10, 37].

The main function of HDL is the transportation of cholesterol and its esters from tissues to the blood circulation and liver for recycling or excretion. Approximately one third of the cholesterol transport is mediated through HDL. Based on this

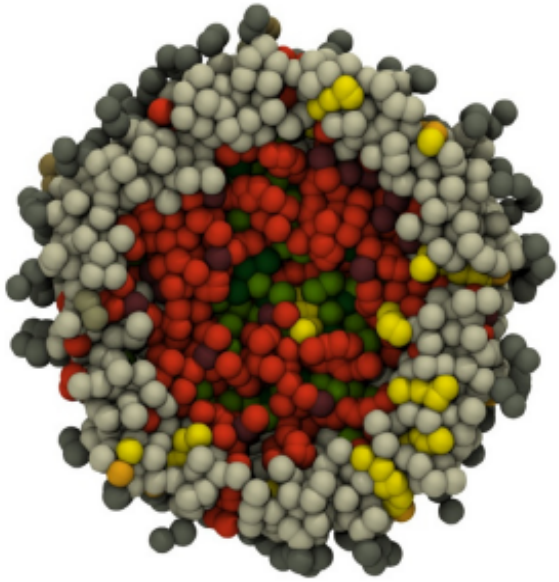


Figure 2.3: Schematic representation of the HDL particle. The molecular components are drawn with different colours: PL headgroups and hydrocarbon chains as dark and light gray, CEs as orange, TG as green and UC as yellow. Modified from [39].

anti-atherogenic function, HDL is often called ”good cholesterol”. The transportation of cholesterol occurs through a process called reverse cholesterol transport (RCT) presented in Figure 2.4 [10]. In general this process involves the acceptance of cellular lipids by HDL, the exchange of lipids between different lipoproteins and HDL and the final delivery of the lipids to the liver for excretion [18, 28]. The size and composition of HDL depends on these different metabolic steps.

When presented in more detail, the process of RCT involves the circulation of lipid-poor HDL particles containing apoA-I and apoA-II. These particles are lipidated with PLs and UC in a process which depends on the adenosine triphosphate (ATP)-binding cassette A1 (ABCA1), see the step one in Figure 2.4 [11, 28]. After lipidation the apolipoproteins form discoidal HDL particles which acquire UC from other lipoprotein particles. The LCAT enzyme esterifies these cholesterol molecules producing CEs when the shape of HDL particles converts from discoidal into spherical as presented in the step two [11, 28]. The composition and shape of HDLs are further modified by CETP which transfers CEs from HDL to LDL and VLDL particles in exchange for TGs. VLDL and LDL then transport CEs to the liver via the LDL-receptor pathway or they are taken up by the liver via the scavenger receptor class B member 1 (SR-B1) pathway presented in the step three [18]. TGs transported to HDL are hydrolyzed by hepatic lipase regenerating small and discoidal HDL particles thus restarting the cycle [41]. According to this, the route of RCT occurs mainly through LDL and

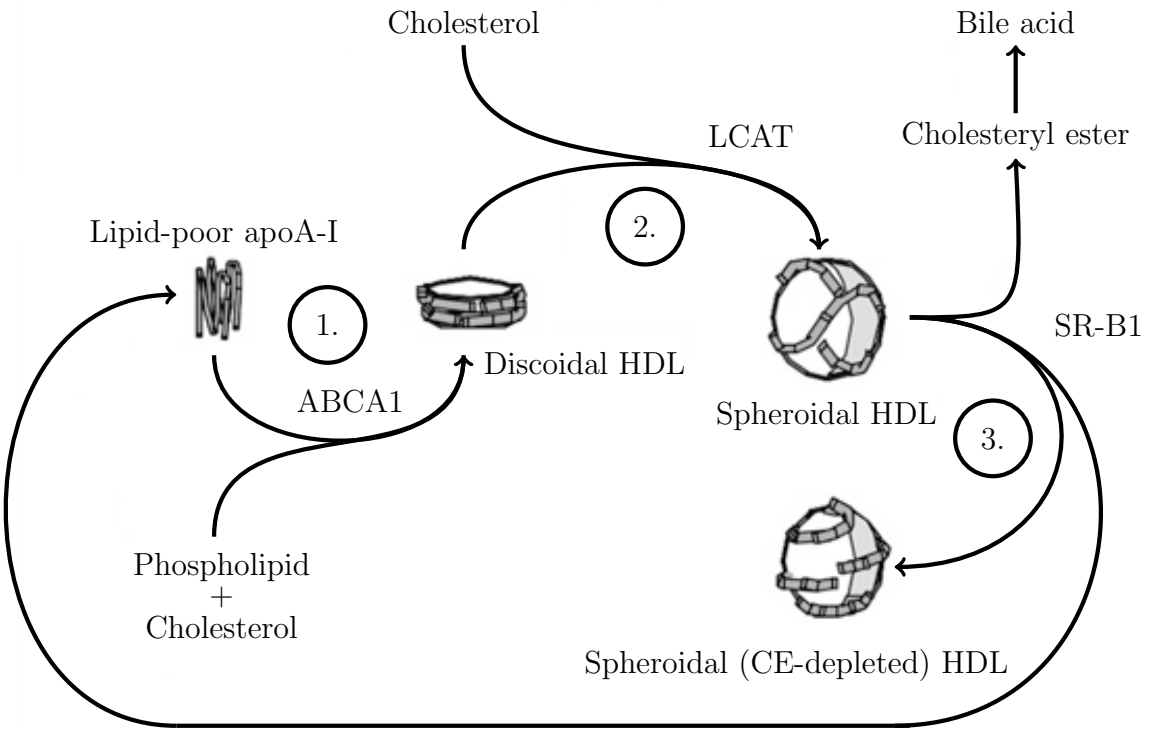


Figure 2.4: Reverse cholesterol transport of cholesterol mediated by HDL. The steps marked as numbers describe the different phases of the process. See text for further details. Modified from [40].

VLDL and partly through HDL. However, if the pro-atherogenic LDL transports cholesterol to tissues more than LCAT is able to esterify, build up of cholesterol begins within the arterial intima [41]. These different metabolic steps have been studied also via MD simulations and the role of simulations has been remarkable during the last decade. Several studies have clarified the molecular structure and functions of HDL involved in RCT. See references [39, 40, 42–46] for further details and discussion.

Besides the essential role of HDL in RCT, it has been proposed to have antioxidant, antithrombotic, pro-fibrinolytic and anti-inflammatory properties against ASCVDs [18]. Since LDL particles undergo the oxidative modification within the arterial intima, the capability of HDL to inhibit this modification is potentially anti-atherogenic. The inhibition can be achieved through the antioxidant or antithrombotic protection mechanisms. The antioxidant property is due to the inhibition of the transmigration of monocytes and the cytotoxicity induced by oxidised LDLs [4, 10, 11]. In the antithrombotic protection mechanism HDL transports several platelet-activating factors that eventually break down the oxidized LDLs and neutralize their proinflammatory effects [11, 17, 47]. The activation of C reactive protein (CRP) due to the production of pro-inflammatory cytokines also plays an

important role in the antithrombotic protection mechanism since this protein is a marker of inflammation [2]. The main action of CRP is to activate complement and to counteract infections [48]. In addition, HDL inhibits the inflammatory mediator cytokine through a process involving an HDL-mediated inhibition of endothelial cell sphingosine kinase [10]. This process is called the anti-inflammatory mechanism and it is influenced by the PL composition of HDL [10]. However, the exact inhibitory mechanisms of these diverse biological mechanisms to atheroprotective effects of HDL remain uncertain, thus further investigations are needed.

As mentioned earlier, during the last decades the research in order to reduce ASCVDs has concentrated on the reduction of the level of LDL-C. Despite the development of numerous useful drugs, ASCVDs continue to be the leading cause of death and disease worldwide and therefore the goal is to find and develop even better treatment methods. The level of HDL-C has been established to be an independent risk factor for these diseases and therefore the studies concentrate more and more on HDL. In addition to the mechanisms of fibrates and niacin, inhibition of CETP has been identified to increase the HDL-C level in the blood. Hence one strategy to reduce ASCVDs is to develop novel molecular agents that inhibit the lipid transfer functions of this protein.

2.2 Cholesteryl ester transfer protein

The connection between the role of CETP and the level of HDL-C was noticed for the first time in Japan in 1989 where in several cases a human genetic deficiency of CETP was related to elevated levels of HDL-C [49, 50]. These studies indicate the inhibition of CETP to be anti-atherogenic. However, some mutation phenotypes suggested increased ASCVD risk based on the ability of CETP to transport CEs from antiatherogenic HDL to proatheronic VLDL and LDL in exchange for TGs [51]. This leads to TG enrichment in HDL as well as to enhanced deposition of cholesterol on the arterial wall [52, 53]. TG enrichment is harmful to HDL since it results in a loss of apoA-I and reduction in size leading to increased catabolism of HDL [53]. Hence CETP can theoretically be considered either antiatherogenic or proatherogenic. This complex relationship gave rise to studies to define the regions of CETP that are important in function through the production of monoclonal antibodies to inhibit the lipid transportation mechanism [54, 55]. In addition, several animal models and early clinical trials have indicated the proatherogenic role of CETP, thus supporting the view that inhibition is a promising strategy for achieving higher levels of HDL-C and lower the risk of severe cardiovascular events [2, 18].

2.2.1 Molecular structure

The human CETP gene exists as a single copy in chromosome 16 in the 16q12-21 region [56]. The gene is expressed at high levels in adipose tissue and the liver [36]. The crystal structure of CETP has been solved by Qiu et al. [55]. Based on that model, CETP is a 472-residue long hydrophobic glycoprotein shaped like a boomerang with dimensions of 13.5 nm x 3 nm x 3.5 nm, see Figure 2.5. The structure can be divided into four units: one terminal at each end of the protein (terminals N and C), a central β-sheet as a linker between the two terminals and a distorted amphipathic helix named helix X, containing the amino acids Glu461-Ser472 at the C terminal domain [55]. Each terminal contains a twisted β-sheet and two helices. In the figure, the helices are denoted A and B in terminal N, and A' and B' in terminal C, with helices B and B' being longer than A and A'. The central β-sheet includes six antiparallel strands consisting of residues before and after the terminals.

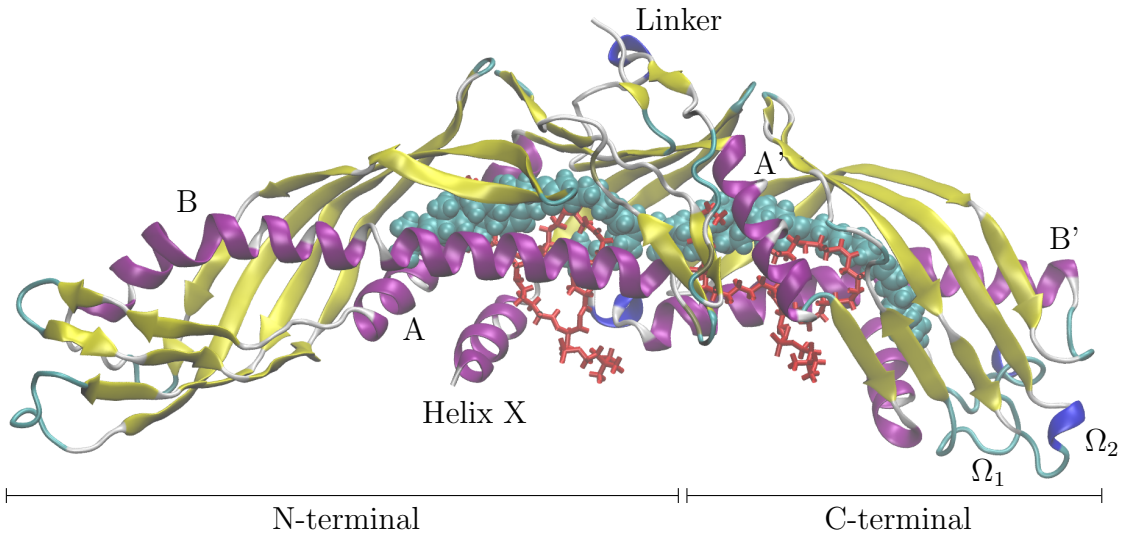


Figure 2.5: Crystal structure of human CETP. The four structural units shown are the terminals N and C, central β-sheet as a linker between the terminals and helix X. Helix X belongs to the C-terminal domain but interacts with residues of the N-terminal domain. Helices A, A', B and B' as well as the loop regions Ω_1 and Ω_2 are labeled. CEs are drawn as cyan space fills of cholesteryl oleate molecules and PLs as red bonds. Modified from [57].

CETP has a 6-nm-long lipid filled tunnel that traverses the core of the molecule and has two distinct openings. The openings are plugged by PLs called dioleoylphosphatidylcholine (DOPC) molecules presented as red bonds in Figure 2.5. DOPCs ($DOPC_1$ and $DOPC_2$) bury their acyl chains inside the hydrophobic tunnel and expose the zwitterionic head groups to solvent [55]. As the binding mode of DOPCs is like this, they shelter the hydrophobic moieties and allow the protein to be more compatible with the aqueous environment. CETP has two neutral lipid binding sites

located inside the tunnel for binding either two CEs, one CE and one TG or two TGs [55]. The figure presents the protein carrying two CE molecules (CE_1 and CE_2) drawn as cyan space fills buried in the middle of the tunnel. CE_1 is located between the terminal N and the central β-sheet whereas the site for CE_2 resides between the central β-sheet and terminal C. Therefore the lipids fill the length of the tunnel and bind in the same orientation with the head of CE_2 packed against the tail of CE_1. The tunnel is formed by a wall of β-sheets underneath the bound lipids and a layer of helices above the lipids [55]. The cross-sectional areas of both openings and the dimensions of the tunnel are large enough to allow the passage of a neutral lipid.

The helix X is located at the C-terminal domain but the hydrophobic face interacts with $DOPC_1$ of the N-terminal domain forming an apolar path to neutral lipid access to the tunnel [55]. The increased B-factor of helix X suggests that the structure of the helix is mobile, Hisb462-Leu464 being a potential hinge region for movements [55]. The B-factor is also known as the atomic displacement parameter and in the protein crystal structure it indicates the fluctuation of an atom about its average position [58]. Therefore, the higher the B-factor, the higher the mobility of an atom in question. Another regions of increased B-factors include the loop regions marked as Ω_1 (residues 284-316) and Ω_2 (residues 346-356) in Figure 2.5 [55]. Helix X has been proposed to act as a lid conducting the neutral lipid exchange by alternating its open and closed states [55, 59]. This is discussed in detail in the next section.

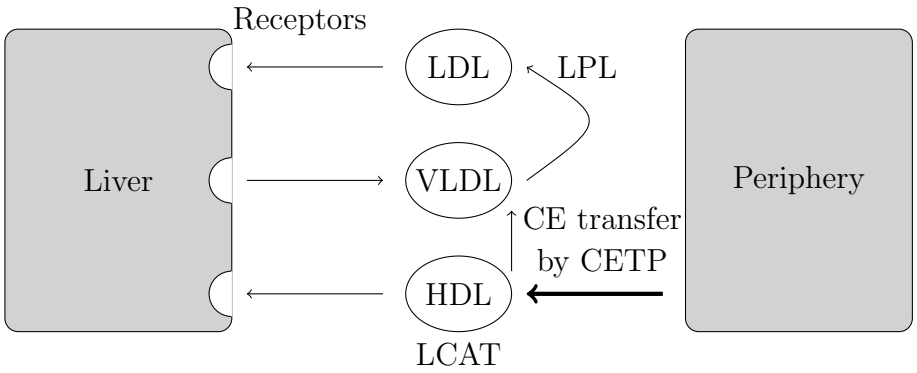
The structure of the concave surface of CETP has been found to be able to bind to lipoprotein surfaces with different curvatures with the modest movements of helix X and the loop region Ω_1 [55, 60]. Other surfaces of CETP do not have the proper curvature to bind spherical lipoproteins and they lack the direct access to the tunnel thus supporting the concept that the concave surface is the only one able to bind lipoproteins. The inherent curvature of CETP makes it well suited to bind HDL-sized particles through its charged and tryptophan residues, indicating the importance of electrostatic interactions during binding [59, 61]. In order to match the curvatures of larger lipoproteins, conformational changes in the structure of the protein occurs. It has been shown that the major part of CETP is associated with HDLs since the affinity of CETP to bind HDL is higher compared with plasma LDL or VLDL [5]. The formation of the CETP-lipoprotein complex is suggested to be modulated by pH, surface pressure and the introduction of positive divalent ions such as Ca^{2+} and Mn^{2+} into the solution [61, 62].

2.2.2 Function

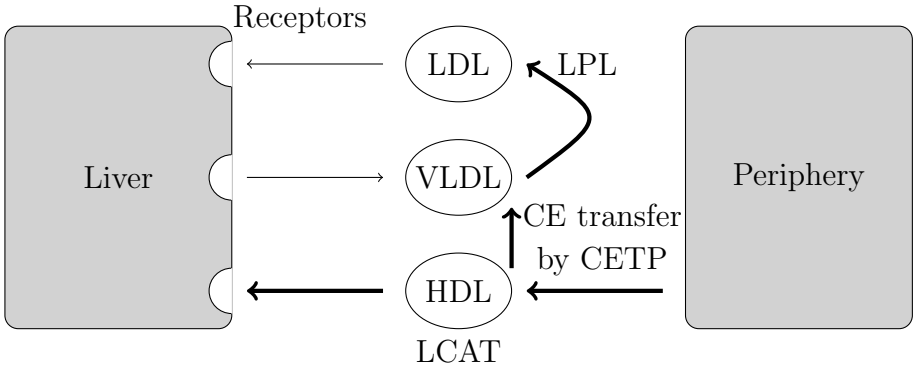
The main function of CETP is to transport and exchange neutral lipids between HDL and other lipoprotein fractions. CETP transports CEs from HDL to VLDL and LDL and TGs in the opposite direction [55]. In addition, CETP transports

also PLs between these particles [18]. Two different mechanisms for the transfer of lipids have been proposed [63]. The ternary complex model suggests that the transfer occurs within a temporary complex between CETP and two lipoproteins [63]. However, the concave surface of CETP indicates that it can bind only one lipoprotein at a time, thus suggesting that the protein operates by a carrier mechanism. First a neutral lipid is accepted from a donor particle to the hydrophobic tunnel, then it is transported through the aqueous phase and finally delivered to an acceptor lipoprotein. The carrier mechanism can occur either through a homoexchange or a heteroexchange mechanism. In homoexchange, CETP applies bidirectional transfer of the same neutral lipid whereas in the more physiologically relevant heteroexchange it transfers different lipids between the lipoprotein particles. The heteroexchange suggests, that when a neutral lipid is accepted by the acceptor particle, CETP cannot detach from the surface of a lipoprotein without acquiring another neutral lipid [60]. In addition, PLs are needed to refill the tunnel openings in order for CETP to be compatible with the aqueous environment [55]. Koivuniemi et al. [59] found in their study that the structure of CETP is unstable without the interior lipids since the hydrophobic tunnel collapsed when CEs were removed from the structure. This finding suggests that the tunnel is not empty at any point during the lipid exchange. CETP mediates the homoexchange mechanism for PLs although the net exchange of these molecules is mainly carried out by the PLTP [55].

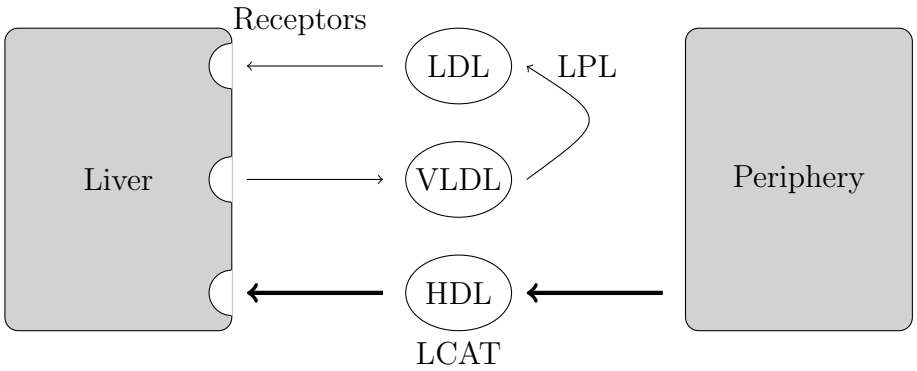
The heteroexchange transportation mechanism is part of RCT and therefore CETP plays a central role in lipoprotein metabolism [64]. The route of RCT and the activity of CETP related to this process are presented in Figure 2.6 including the presence of normal as well as increased activities and the absence of CETP. The initial step is the transportation of dietary cholesterol to peripheral tissues through LDL. Excessive cholesterol is removed by HDL with the help of LCAT, which provides the driving force for the removal of cholesterol. The rate of lipid transfer between different lipoprotein particles is influenced by an increased TG level, eventually resulting CE depletion and TG enrichment in HDL [65]. Combined activities of CETP and hepatic lipase induce the formation of smaller dense HDL particles that are more rapidly catabolised [5, 65]. Due to this, the route of RCT occurs mainly via LDL and VLDL which eventually are taken up by the liver via SR-B1. As long as this process to remove LDL and VLDL stays normal, these particles do not accumulate excessively into the intima. However, during the increased activity of CETP the rate of CE transportation to LDL and VLDL is increased leading to a reduction in the size of HDLs as well as to their increased catabolism [53]. Consequently, CE-enriched LDL particles accumulate in the intima. Increased CETP activity is caused by increased CETP gene expression in peripheral tissues or increased concentrations of acceptor lipoproteins [5].



(a) Normal CETP activity



(b) Increased CETP activity



(c) CETP deficiency

Figure 2.6: Models of RCT and the related activity of CETP in three different metabolic settings. a) RCT in the presence of normal CETP activity. b) RCT when the activity of CETP is increased. c) RCT in the absence of CETP in humans with genetic CETP deficiency. See text for further details. Drawn using [5] as a reference.

During the deficiency of CETP, all its neutral lipid transfer functions are absent. Therefore the content of TGs is decreased and the content of CEs is increased in HDL particles. Because of this and the continued activity of LCAT, HDL particles gradually increase in size and acquire apolipoprotein E, thus becoming a high affinity ligand for the LDL receptors of the liver. These receptors eventually remove the increased HDL particles and, moreover, also the cholesterol from tissues [5]. The absence of CETP results in a build up of HDL to very high levels and thus the amount of cholesterol transported in VLDL and LDL is reduced [66]. Due to these facts the deficiency of CETP correlates with reduced risk of ASCVDs. The mechanism of CETP inhibition to raise the HDL-C level is suggested to occur through the stimulation of cholesterol efflux and anti-inflammatory effects in macrophage foam cells [67]. Therefore the route of RCT is changed to occur through HDL instead of VLDL and LDL.

It has been proposed that during lipoprotein binding the PLs plugging the openings of the hydrophobic tunnel may merge into the PL monolayer of lipoprotein to permit the access of neutral lipids into the tunnel [55]. In addition, helix X has been suggested to move aside from the N-terminal opening, thus facilitating the lipid exchange [60]. Based on the MD simulations performed by Koivuniemi et al. [59], it was shown that during binding the PLs migrated away from the tunnel openings forming a small hydrophobic patch under the concave surface of CETP. In addition, these simulations showed that the conformation of helix X rearranged and became buried inside the hydrophobic tunnel instead of the PL monolayer [59]. According to these observations helix X is claimed to play a central role in the process of neutral lipid exchange. It is suggested to act as a lid, conducting the lipid exchange, being locked to the open state for the time of exchange and to the closed state during lipid transfer. Therefore there are two important functional properties of the helix that make the lipid exchange possible: management of the lipid exchange by opening the hydrophobic pathway from a lipoprotein surface to the hydrophobic tunnel and the improvement of the diffusion of neutral lipids from the hydrophobic tunnel to lipoproteins by filling the volume of CETP-bound neutral lipid when it diffuses out from CETP [59].

According to the given suggestions, the movement of helix X aside from the N-terminal tunnel opening is considered to be sufficient for the lipid exchange to occur. However, this is in contradiction with several mutational studies that showed that when helix X was completely removed from the structure of CETP, the protein was able to exchange only PLs but not CEs or TGs [68, 69]. For this reason the role of helix X may not be as simple as previously suggested. Instead, it is probable that it interacts according to a currently unknown mechanism with CEs, thus facilitating their exchange, but the confirmation of this issue requires further research.

In any case, the conformational changes of helix X can be considered to have an important role in the inhibition of CETP. One possible strategy to achieve the inhibition is the development of nonpolar drugs that are then transferred into the hydrophobic tunnel of CETP. Based on MD simulations done by Koivuniemi et al. [59] it can be assumed that helix X is locked to the open state when inhibitors are bound to CETP. This may be due to the fact that the inhibitor changes the conformational distribution of helix X to favor the open state. This would stabilize the CETP-HDL complex since CETP would not be able to detach from the surface of HDL without helix X to shield the hydrophobic tunnel opening. In addition, CETP may not be able to bind and transport PLs since either the open state or the drugs could prevent the binding of PLs to the C-terminal opening thus further stabilizing the complex.

2.3 Inhibitors of cholesteryl ester transfer protein

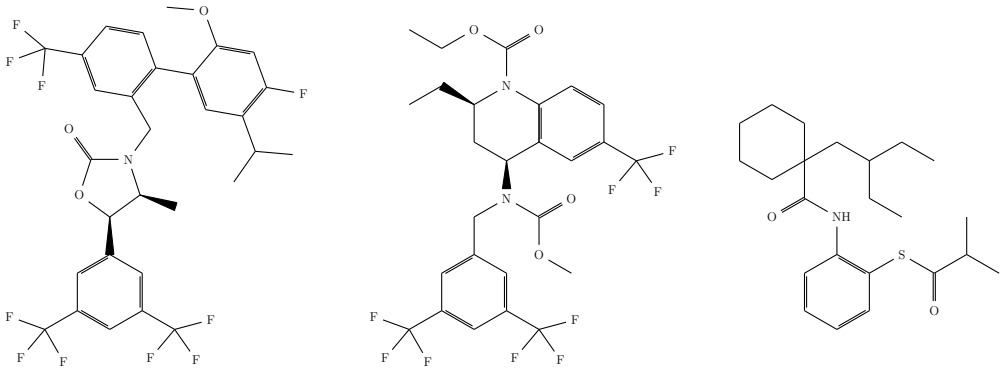
The most promising strategy up to date to increase the HDL-C level is to inhibit the lipid transfer functions of CETP. This can be achieved through the use of novel molecular agents which, depending on the inhibitory mechanism, prevent either the binding of neutral lipids or their release once bound or initiate the excessive or insufficient binding of CETP to lipoproteins. As these diverse inhibitory actions suggest, the precise mechanism requires further elucidation. This work attempts to clarify this issue behind the novel agent anacetrapib, which has been shown to effectively raise the HDL-C level with an acceptable side effect profile in patients with ASCVDs or risk factors for ASCVDs [7]. Understanding the functions of CETP as well as the mechanisms behind CETP inhibitors are important to realize in order to develop safe treatment methods for the pharmacological raising of HDL-C.

2.3.1 Anacetrapib

Structure-activity relationship

Anacetrapib, previously denoted as MK-0859, is an orally active, potent and selective CETP inhibitor developed to treat ASCVDs. It was identified by high-throughput screening to belong to the 1-3-oxazolidin-2-one series of CETP inhibitors developed by Merck [70]. The chemical name of the agent is denoted as (4S,5R)-5-[3,5-bis(trifluoromethyl)phenyl]-3-[4'-fluoro-2'-methoxy-5'-(propan-2-yl)-4-(trifluoromethyl)[1,1'-biphenyl]-2-yl]methyl-4-methyl-1, 3-oxazolidin-2-one and the molecular formula is $C_{30}H_{25}F_{10}NO_3$ [71]. The structural formulas of all the current CETP inhibitors anacetrapib, torcetrapib and dalcetrapib are presented in Figure 2.7.

As can be seen from Figure 2.7, anacetrapib shares some structural similarities



(a) Anacetrapib (b) Torcetrapib (c) Dalcetrapib

Figure 2.7: Structural formulas of different CETP inhibitors. Drawn using [71] as a reference.

with torcetrapib but is different from dalcetrapib. Anacetrapib consists of three aromatic biphenyl (C_6H_5) rings connected with a 2-oxazolidinone ring ($C_3H_5NO_2$). In addition, it has three trifluoromethyl (CF_3) groups as well as three methylene (CH_3) groups connected to its biphenyl rings. The biggest differences between anacetrapib and the other two inhibitors are the absence of the 2-oxazolidinone group from both torcetrapib and dalcetrapib as well as the absence of trifluoromethyl groups from dalcetrapib. It has been reported that modifications of the oxazolidinone ring led to the discovery of anacetrapib [72]. The medicinal chemistry behind this discovery is described in [72].

Both anacetrapib and torcetrapib have been reported to increase the binding affinity of CETP towards lipoproteins and to induce a tight reversible binding making the HDL-CETP complex more stable [70, 73]. This mechanism of action differs from the one performed by dalcetrapib which forms an irreversible disulfide bond with the cysteine at residue 13 of CETP [74]. The increased binding has been shown to inhibit the CETP-mediated cholesterol exchange between lipoproteins. The potency of anacetrapib to inhibit the transfer of CEs described with the half maximal inhibitory concentration (IC_{50}) has been shown to be 17 ± 4.8 nM with similar values for the transfer of TGs [75]. Torcetrapib has the similar inhibitory potencies of CEs and TGs compared with anacetrapib, whereas dalcetrapib has a 70- to 80-fold weaker potency ($IC_{50} = 1178 \pm 443$ nM) [75]. This indicates that anacetrapib is at the moment the most promising agent to increase HDL-C.

The activity of anacetrapib has been studied using pharmacokinetics and pharmacodynamics. This is performed by investigating the absorption, distribution, metabolism and excretion profiles of anacetrapib in a single oral dose study on humans [76]. The results of this *in vitro* study suggested that anacetrapib probably applies a low-to-moderate degree of oral absorption later supported by the pharma-

cokinetics study of anacetrapib in rats and rhesus monkeys [77]. It was found out that the absorbed fraction of the dose is metabolized mainly via CYP3A4-catalyzed oxidative metabolism followed by the excretion of three oxidative metabolites through the biliary-fecal route [76]. CYP3A is an enzyme common to the metabolism of a number of drugs, e.g. statins, used in the treatment of ASCVDs [78]. The three metabolites of anacetrapib excreted during the process of metabolism involved *O*-demethylation which was further hydroxylated at the biphenyl and isopropyl moieties [76, 77]. The finding concerning the metabolism and clearance of anacetrapib was further supported by the clinical drug-drug interaction study, which investigated whether anacetrapib is an inducer or an inhibitor for the CYP3A activity. The study showed that catabolism of anacetrapib does not seem to interfere with other drugs metabolized through the CYP3A4 pathway [78]. In addition, anacetrapib was found to stay intact after metabolism and due to these facts, the oxidative metabolism can be considered as the major route of eliminating anacetrapib from the blood circulation [76]. In general anacetrapib is a well-tolerated agent but the long-term safety and tolerability of the drug remains to be evaluated.

The corresponding profiles of torcetrapib were compared with the profiles of anacetrapib. The comparison revealed significant differences between the metabolism and the routes of elimination of these two inhibitors. A total of 15 primary and secondary metabolites of torcetrapib were identified in humans and these were eliminated mainly through urinary excretion [76]. This difference is suggested to be related with the differences in the metabolic pathways of the agents. As anacetrapib applies the oxidative metabolism, the key feature in the metabolism of torcetrapib was reported to be related with an oxidative cleavage [76]. This induces the formation of low-molecular weight metabolites including quinaldic acid and bistrifluoromethylbenzoic acid which were excreted via the renal pathway [76].

Clinical trials

The ability of anacetrapib to increase the level of HDL-C as well as to decrease the level of LDL-C related to the inhibition of the lipid transfer functions of CETP has been studied in phase one, two and three clinical trials performed in humans. The early phase one trial involved patients with dyslipidaemia as well as healthy individuals receiving either anacetrapib or placebo. The effect of anacetrapib on lipoprotein levels was studied from the dyslipidaemic patients and the effect on blood pressure was examined from the healthy participants. The results were promising since a 129 % increase in HDL-C and a 38 % decrease in LDL-C were documented in patients with dyslipidaemia and no effect on blood pressure was observed with healthy individuals [79]. Similar increases in lipoprotein cholesterol levels were observed in

a phase two study where 589 patients with dyslipidaemia were randomly assigned to anacetrapib. Results showed up to 139 % increase in HDL-C and up to 40 % decrease in LDL-C without an effect on blood pressure [80].

Determining the Efficacy and Tolerability of CETP Inhibition With Anacetrapib (DEFINE)-trial recruited 1623 patients with coronary heart disease or at high risk for coronary heart disease already receiving statin therapy for an anacetrapib phase three trial [81]. The patients were divided into two groups from which one received anacetrapib and the other a matching placebo. After one and a half years, the level of LDL-C had reduced from 2.1 mmol/dL to 1.2 mmol/dL in the anacetrapib group whereas the corresponding numbers in the placebo group were decreased from 2.1 mmol/dL to 2.0 mmol/dL. This means a 39.8 % reduction in the anacetrapib group when compared with the placebo group. Correspondingly, the level of HDL-C had increased from 1.0 mmol/dL to 2.6 mmol/dL in the anacetrapib group, compared with an increase from 1.0 mmol/dL to 1.2 mmol/dL in the placebo group - a 138.1 % greater increase in the anacetrapib group. Therefore the cholesterol values obtained in the anacetrapib group corresponded well with the general guidelines presented in Section 2.1.1. The apoB level decreased by 21 % and apoA-I level increased by 44.7 % in the anacetrapib group beyond the changes seen in the placebo group, thus indicating that the overall effects of anacetrapib on the cholesterol levels are two to four times as large as those with other CETP inhibitors tested until now. Anacetrapib was also stated to have an acceptable side-effect profile since changes in the blood pressure or at the levels of aldosterone and electrolytes were not observed in comparison with torcetrapib.

Despite the promising results, concerns have been raised whether the larger, cholesterol-rich HDL particles generated through CETP inhibition remain antiatherogenic. These concerns have been defused by the investigations of HDL particles isolated from patients treated with torcetrapib and anacetrapib. It has been shown that these HDL particles promote even enhanced RCT and cholesterol efflux from macrophages [82, 83] through the ATP-binding cassette subfamily G (ABCG)-transporter [11, 67]. The higher the level of HDL-C, the more enhanced these functions are.

2.3.2 Other inhibitors

Torcetrapib

The first developed CETP inhibitor was torcetrapib (4-[(3,5-Bis-trifluoromethylbenzyl)-methoxycarbonyl-amino]-2-ethyl-6-trifluoromethyl-3,4-dihydro-2H-quinoline-1- carboxylic acid ethyl ester). In the early small phase one studies torcetrapib was shown to increase the level of HDL-C from 6 % to 91 % and to decrease the level of

LDL-C from 21 % to 42 % in healthy subjects and patients with HDL-C less than 2.2 mmol/dL [84, 85]. In addition, these trials showed that the level of apoA-I was elevated and the level of apoB was reduced.

Torcetrapib was evaluated on a much larger scale through the Investigation of Lipid Level Management to Understand its Impact in Atherosclerotic Events (ILLUMINATE)-trial. 15067 patients with a history of ASCVD were assigned to this randomized, double-blind trial [6]. The patients were divided into two groups, one received torcetrapib plus atorvastatin whereas the other received only atorvastatin. During the 18-month-study the patients receiving torcetrapib experienced a 72.1 % increase in HDL-C and a 24.9 % decrease in LDL-C [6]. The changes in the cholesterol levels were shown to be caused by the ability of torcetrapib to promote the formation of CETP-lipoprotein complex thus blocking both neutral and PL transfer activities of CETP [66]. However, ILLUMINATE trial was prematurely terminated because of an increase in mortality and morbidity caused by cancers and infections. Torcetrapib was shown to increase the blood pressure of 3-6 mmHg, decrease the serum potassium level and increase the levels of serum sodium, bicarbonate and aldosterone [6]. The possible causes for the failure of torcetrapib are these off-target effects together with an adverse effect of CETP inhibition in the generation of dysfunctional HDL particles not participating in RCT [86, 87].

The effects of torcetrapib on the progression of coronary atherosclerosis, carotid atherosclerosis in familial hypercholesterolemia and carotid intima-media thickness in mixed dyslipidaemia were studied in The Investigation of Lipid Level Management Using Coronary Ultrasound to Assess Reduction of Atherosclerosis by CETP Inhibition and HDL Elevation (ILLUSTRATE) [86], in the Rating Atherosclerotic Disease Change by Imaging With a New CETP Inhibitor 1 (RADIANCE 1) [88] and in the RADIANCE 2 [87] trials, respectively. Despite the elevated HDL-C level and reduced LDL-C level observed in all these trials, torcetrapib was not stated to reduce the progression of any of these diseases. In addition, in every trial torcetrapib was associated with an increase in blood pressure similar with the ILLUMINATE trial. The relationship of these lipid changes and progression of coronary atherosclerosis was studied in the post hoc analysis and it was observed that the patients with the greatest HDL-C increase did have some regression of the disease caused by the modest increase in the efflux of cholesterol from macrophages into HDL [82, 89]. Despite these controversial results, it may be concluded that the generated HDL particles are functional and participate in RCT. This indicates that the concept of CETP inhibition to protect against ASCVDs remains viable and thus the failure of torcetrapib was related to the molecule rather than the mechanism of action.

The reasons behind the lethal side effects caused by torcetrapib has been explained through the interactions of the drug with unknown off-targets since CETP inhibitors

have been suggested to bind several proteins leading to the observed side effects [90]. This protein-ligand network can be identified using the approach denoted as chemical systems biology. The approach systematically explores the interactions between proteins and the ligand incorporating them into biological pathways [90]. This provides valuable information about the molecular mechanisms behind the side-effect profile of torcetrapib. The identified off-targets were reported to belong to different protein families but they are structurally and functionally related to CETP [90]. These proteins are mainly involved in the lipid metabolism, immune response and signaling networks. It is suggested that the number of side-effects caused by torcetrapib can be minimized by fine tuning the multiple off-target interactions by introducing molecules that bind to those molecular components that are involved in the regulation of negative metabolism [90].

Dalcetrapib

Dalcetrapib, also known as JTT-705, is a thioester S-(2-((1-(2-ethylbutyl)cyclohexane) carbonylamino)phenyl)2-methylpropanethioate [91]. It is the only competing inhibitor for anacetrapib currently on the market. Dalcetrapib has been reported to bind covalently to CETP by forming a disulphide bond with cysteine 13 locating inside the hydrophobic tunnel of CETP [74]. This bond formation is suggested to be the essential requirement for the dalcetrapib-mediated CETP inhibition since it induces a conformational change in the structure of CETP, which could increase the stability of HDL-CETP complex [65, 75]. Therefore, in addition to the binding mode, differences exist also in the inhibitory mechanism when comparing dalcetrapib with torcetrapib and anacetrapib. However, more investigations are needed in order to better compare the different inhibitors with each other.

In the early study dalcetrapib was shown to increase the level of HDL-C and to inhibit the progression of atherosclerosis in cholesterol-fed rabbits [74]. However, in rabbits with more severe hypercholesterolemia, dalcetrapib did not reduce the risk of atherosclerosis despite increases in HDL-C [92]. Despite these controversial results, dalcetrapib increased HDL-C up to 34 % and decreased the activity of CETP by 37 % in healthy patients with mild hyperlipidemia in a randomized phase two study [93]. Another phase two study involving patients with dyslipideamia on pravastatin therapy showed 28 % increase in HDL-C and a 30 % decrease in CETP activity [94]. Neither of these studies showed an increase in blood pressure or adverse cardiovascular events, as seen in the trials concerning torcetrapib. These findings were supported in the safety and tolerability trial of dalcetrapib [95]. Currently a phase three clinical trial with 15600 coronary syndrome-patients concerning dalcetrapib's effects on cardiovascular morbidity and mortality is ongoing [96]. It is estimated to be completed in 2012 [92].

3. MOLECULAR DYNAMICS

In this thesis, the interactions between CETP and anacetrapib are studied with classical MD simulations. These simulations are a computer aided way to calculate the movements and interactions of systems consisting of several small particles using classical mechanics. Compared with experimental measurements, the MD simulations have many advantages but also disadvantages. Understanding these is important in order to interpret the obtained results accurately and correctly. This chapter gives an overview of the MD simulation method applied in this study. The overview is based mostly on the manual of the applied software, GROMACS [97], since it gives a solid description of MD.

3.1 The classical molecular dynamics method

The real life experiments and MD simulations have many similarities with each other despite the fact that the experiments are performed in a laboratory and the MD simulations on computers. An experiment begins with a sample preparation followed by the actual tests and measurements. In order to reduce the random statistical noise and the possibility of measurement errors more samples are prepared, measured and averaged. In a similar manner, the MD simulation is started by first constructing the system to be simulated with a topology and a force field describing the structures of particles and the interactions between them. After this, the system is equilibrated from an initial configuration consisting of the initial coordinates and velocities of the system. This is done by using Newton's equations of motion. After equilibration, the simulation can be started.

One major disadvantage of experimental measurements is that they are difficult to perform in small scales - in nanoscales this is almost impossible. Hence one solution to study the behavior of small particles is to use MD simulations. They offer a possibility to examine the small scales with detailed knowledge on an atomic scale. The second significant advantage of MD simulations is their cheap implementation. The primary disadvantages are the inaccuracy and the small sizes of the systems to be modeled. Both of these problems are mainly caused by the large computational need required by the totally atomistic MD simulations. Even though the modern calculation units consist of several thousands of processor cores, models used to simplify the description of interactions and movements of systems over irrelevant details must be developed.

With the help of these models it is possible to approximate and combine different things, and in this way the computation time will be considerably shortened.

Additionally also the time scales of biological processes compared with the time scales of simulations are a source of problems. Essentially the problem is caused by the fact that the modeling methods currently available are not able to operate both reliably and efficiently over long time scales. As an example, the so called *ab initio*-methods allow a very detailed examination of a particle over a scale smaller than the atomic one. These methods take the quantum-mechanical nature of the particle into consideration, but the calculations are computationally very heavy and the studied systems very small. A good example is the folding problem of proteins with different sizes. The folding of small proteins from the open structure towards the equilibrium state takes at least one microsecond. By using the *ab initio*-methods, these transformations can be studied only up to a few picoseconds and a single computer requires approximately one year for the calculation. Hence in order to understand the comprehensive operation of a biological system, other methods have to be considered. One solution is to use classical mechanics which neglects the quantum mechanics. The time needed in calculations is shorter and hence the systems can be larger. Summarizing, the understanding of biological systems requires the exploitation of all the different modeling approaches.

As mentioned, the MD simulation begins with creating an initial structure for the system to be simulated and giving the required parameters for the simulation. Every simulated particle is simply thought of as a center of mass to which the interactions are directed from the force field and the potential functions related to it. The core of MD simulation is a loop where the force field is used to calculate the forces interacting between the particles, and according to these, the positions and velocities of each particle are updated using Newton's equations of motion. The loop is repeated with a regular interval until enough simulation data has been gathered. In addition, the simulation is assumed to follow the ergodic hypothesis meaning that the time average of the system is equivalent to the ensemble average of the system. However, before the MD simulations could have been invented and utilized in different studies, investigations concerning the molecules forming the biological world and their structures have been carried out. The formation of a constantly developing and freely available molecule databank provides wide opportunities for biomolecular modeling.

3.2 Structure of the molecule

The field of biomolecular modeling dates back to the 1960s and nowadays it is developing rapidly. Advances are driven by improvements in computing power, memory capabilities, algorithms for molecular dynamics as well as the developing

and versatile genomic and structural databases. The first records of crystallized polypeptides and proteins date to the end of the 1920's and the beginning of the 1930's [98]. The techniques to clarify the crystal structures more carefully developed further and in late 1950's the first x-ray diffraction pattern from the crystal structure of a protein was obtained. Structure determination based on x-ray crystallography involves the analysis and comparison of the x-ray diffraction pattern of a native protein to those associated with the protein bound to heavy atoms like mercury [98]. Another important technique for solving biomolecular structures is nuclear magnetic resonance spectroscopy. With the help of this method it is possible to obtain information about the interactions and the localized motions of the molecules [98].

With the aid of protein crystallization techniques it has been possible to relate the structural information of a system to its functions. It is known that proteins play crucial roles as constituent molecules and as triggers of physiological processes for all living things. Proteins for instance provide the support in muscle tissues, ligaments, tendons, bones and organs. They also play crucial regulatory roles in transport and storage, such as oxygen in muscle and blood cells. Moreover, the proteins make the moving, signal transmission and immuno defence of cells possible. Protein itself is an organic compound consisting of amino acids linked together as a chain. Amino acids bond themselves to each other with peptide (N-C) bonds. Each amino acid consists of a central tetrahedral carbon known as the alpha carbon (C^{α}) attached to a hydrogen atom, a protonated amino group (NH_3^+), a dissociated carboxyl group (COO^-) and a varying side chain [98]. There are 20 naturally occuring amino acids of which humans can synthesize about a dozen thus leaving the remaining amino acids to be ingested through the diet.

The structure of a protein can be classified with four basic levels: primary structure (the sequence of amino acids), secondary structure (regular local structural patterns such as α-helices and β-sheets), tertiary structure (the three-dimensional arrangement of all atoms in the protein in space) and quarternary structure (the complete 3D interaction network) [98]. Most proteins adopt a specific 3D conformation which is related to their biological activities. The conformations of proteins are composed of four secondary structural elements, namely the α-helices, β-sheets, turns and loops. The classical right-handed α-helix connects each backbone carbonyl (C=O) oxygen of residue i to the backbone hydrogen of the NH-group of residue $i + 4$ and therefore the helix has 3.6 residues per turn [98]. The helix may be curved depending on the amino acid sequence. 3_{10}- and π-helices are common variants of the α-helix motif and they are usually unstable in solution. 3_{10}-helix is tighter compared with the α-helix since it has 3 residues per turn involving hydrogen bonds between residues i and $i + 3$ [98]. The nomenclature reflects that there are 10 atoms within the hydrogen bond. π-helix in turn is more loosely coiled having hydrogen bonds

between residues i and $i + 5$ [98]. Other helix types are left-handed α-helix and collagen helix. The second basic element is the β-sheet which forms by aggregating amino acid strands (β-strands) via hydrogen bonds [98]. β-strands are 5-10 residues long and the aggregation can occur in a parallel or anti-parallel orientation. Turns can be found in regions where the sharp reversals of orientation reside, such as in the junction between two parallel β-strands whereas loops occur in short regions connecting various motifs [98]. The majority of these elements lie on the surface of the protein because of the solvation considerations, since most loops are highly hydrophilic and interact with solvent [98].

The structure of a biomolecule, as obtained from the x-ray crystallography, is still insufficient for the proper understanding of its functions and biological activity, since it only provides a frozen view of a complex system. In addition, attention should be paid to how well the x-ray structure describes the protein in its native environment, because the x-ray structure may include packing effects. Molecules are members of living matter whose constituent atoms continuously interact with each other as well as with their surroundings. At this point the MD simulations join the picture since the molecular conformation, motion and the associated biological functions of different molecules can be studied with the help of simulations. In this way it is possible to connect the structure of a complex system to its functions.

3.3 Topology

Every MD simulation requires a molecular description of the particles to be simulated. The description is called a topology and it contains the size and shape as well as the complete description of all the interactions that occur within and between the particles in the simulated system [99]. One commonly applied method - and the one applied in this study - is to produce a topology for a system, which includes every atom in the description. If the simulations need to be faster, it is possible to use a united-atom representation where the number of atoms is decreased by representing methyl and methylene groups as a single particle [99]. The process to simplify the description is called coarse graining but it is not discussed here further.

With the help of molecule databanks providing information about experimentally-determined structures of proteins, nucleic acids and other complex assemblies, it is possible to easily achieve the topology of a molecule. However, when there does not exist a ready topology, it has to be constructed. The construction is not always a straightforward process since it may require a manual placing of particles to their desired positions, and afterwards the energy minimization of this artificial structure to remove the steric clashes and inappropriate positions of atoms. How this is done depends on the system under study. Neither is the parametrization of a particle straightforward since this process begins with the

considerations of intramolecular potential energies and atomic partial charges through quantum mechanical calculations. These are then held fixed while the non-nonded interactions are determined. The results are compared with experimental data and the parameters are adjusted if needed.

3.4 Force field

The interactions between different particles are described through a force field. The interactions can be divided into bonded and non-bonded. Bonded interactions describe the interactions between particles connected by a chemical bond, including the stretching of the bond and the bending of the bond angle and dihedral angle. Non-bonded interactions desribe the interactions between any pair of particles including the weak van der Waals-type and electrostatic Coulombic forces. These interactions are pair-additive and centro-symmetric in the MD software (GROMACS) applied in this study. The terms of the mathematical equations describing the interactions are derived from both experimental work and high-level quantum mechanical calculations.

3.4.1 Bonded interactions

The bonded interactions include not only the interactions between two atoms but also the 3- and 4-body interactions. 2-body interactions are denoted as bond stretching, 3-body interactions as bond angle and 4-body interactions as dihedral angle. The bond stretching between two covalently bonded atoms i and j can be described by a harmonic potential

$$V_b(r_{ij}) = \frac{1}{2}k^b_{ij}(r_{ij} - b_{ij})^2, \tag{3.1}$$

where k_{ij} is the force constant describing the rigidity of the bond, r_{ij} is the distance between the particles and b_{ij} is the reference distance.

The bond angle vibration between a triplet of atoms $i - j - k$ is represented with a similar manner as the bond distance vibrations, namely using a harmonic potential on the angle θ_{ijk}:

$$V_a(\theta_{ijk}) = \frac{1}{2}k^\theta_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2, \tag{3.2}$$

where θ_{ijk} is the angle between the three atoms, k_{ijk} is the force constant and θ^0_{ijk} is the reference angle.

A proper dihedral angle describes the angle between two planes formed by four particles. These angles can either be periodic or of the Ryckaert-Belleman -type and

they are defined according to IUPAC/IUB convention where ϕ is the angle between the planes ijk and jkl. Zero angle corresponds to the *cis* configuration which means that i and l are on the same side. The dihedral interaction can be described as

$$V_d(\phi_{ijkl}) = k_\phi(1 + cos(n\phi_{ijkl} - \phi_s)), \tag{3.3}$$

in which ϕ_{ijkl} is the angle between the four atoms, k_ϕ is the force constant, n is a selectable constant and ϕ_s is the reference angle. Improper dihedral is a special type of dihedral interaction used to force the atoms to remain in a plane or to prevent transition to a mirror image.

3.4.2 Non-bonded interactions

The non-bonded interactions usually contain a repulsion term, a dispersion term and a Coulomb term. The repulsion and dispersion terms are combined in terms of the weak van der Waals interactions described by the Lennard-Jones (LJ) potential:

$$V_{LJ}(\mathbf{r}_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{3.4}$$

where r_{ij} describes the distance between the particles i and j. The parameters ϵ_{ij} and σ_{ij} are chosen separately for each pair of particles. ϵ_{ij} describes the minimum energy of the potential well and σ_{ij} describes the distance when the potential changes from repulsive to attractive.

The classical Coulombic potential describes the electrostatic interactions between two charged particles:

$$V_C(r_{ij}) = \frac{1}{4\pi\epsilon_0\epsilon_r}\frac{q_i q_j}{r_{ij}}, \tag{3.5}$$

where ϵ_0 is the permittivity in vacuum and ϵ_r the relative permittivity. q is the charge of the particle in question.

3.5 Equations of motion

The molecular dynamics approach is simple in principle. We simulate motion of a system under the influence of a specific force field by following molecular configurations in time according to Newton's equations of motion for a system consisting of N interacting atoms with positions $\{\mathbf{r}_i\}$ and masses $\{m_i\}$:

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = \mathbf{F}_i, i = 1...N. \tag{3.6}$$

The total force, $\mathbf{F}_i$, acting on a particle i is the negative derivative of a potential function $V(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$:

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \tag{3.7}$$

These two equations are solved simultaneously in small time steps. The most time-consuming part of a MD simulation is the computation of the forces in the system. In order to minimize the amount of computation, the forces acting on a certain particle are calculated only for a certain group of neighbour atoms that lie within the cut-off radius for the interaction. The information of which interactions to calculate is contained in the neighbour lists.

In order to calculate new coordinates and velocities for the next time step using the above described equations of motion, a time integration algorithm is required. The GROMACS MD program utilizes the so-called leap-frog algorithm which uses coordinates $\mathbf{r}_i$ at time t and velocities $\mathbf{v}_i$ at time $t - \frac{\Delta t}{2}$ in order to update the positions and velocities using the forces $\mathbf{F}(t)$ at time t. The equations of the leap-frog algorithm for particle i are written as follows:

$$\mathbf{v}_i(t + \frac{\Delta t}{2}) = \mathbf{v}_i(t - \frac{\Delta t}{2}) + \frac{\mathbf{F}_i(t)}{m}\Delta t; \tag{3.8}$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{\Delta t}{2})\Delta t. \tag{3.9}$$

The algorithm is of third order in r, time-reversible and equivalent to the Verlet algorithm [100]. See reference [101] for further details and comparison with other time integration algorithms.

3.6 Periodic boundary conditions

Several ways to treat the boundaries - and therefore to minimize the edge effects caused by a small isolated system - of the simulated system exist. The classical way is to use periodic boundary conditions (PBCs) where the atoms of the system to be simulated are put into a space-filling box surrouded by translated copies of itself thus forming an infinite lattice [102]. Therefore, as the particle moves in the original box, its periodic image in each of the neighbouring boxes moves in exactly the same way. Due to this, if the particle crosses one side of the simulation box, it

immediately re-enters from the opposite side. Since the particles interact with its nearest neighbours, it does not matter whether the neighbour is in the original box or in one of its copies because the interactions are calculated over the box boundaries. Attention should be paid to the size of the box since if it is defined to be too small, the atoms residing near the box boundaries will experience the interactions caused by the other atoms more than once and from several directions causing the unwanted edge effects. When the PBCs are applied in the simulations concerning CETP and anacetrapib, the quadratic simulation box containing the protein and the drug is copied on every side of the original box and the same is done for these new boxes. Hence the system can be considered as infinite with an infinite number of proteins and drugs surrounded by water.

3.7 Temperature and pressure coupling

Maintaining the constant temperature through a simulation can be achieved for instance through the Berendsen or the Nosé-Hoover coupling [103, 104]. In the Berendsen coupling, the system is weakly connected to an external heat bath given temperature T_0. The temperature difference decays exponentially with a time constant τ:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \tag{3.10}$$

The advantages of this coupling method are that the strength of the coupling can be varied to adapt the user requirement as well as that the algorithm is extremely efficient to equilibrate the system to the target temperature [103]. However, the Berendsen coupling method suppresses the fluctuations of kinetic energy meaning that a proper canonical ensemble cannot be generated. This can be achieved through the Nosé-Hoover coupling applied in this study. The Nosé-Hoover method introduces a thermal reservoir and a friction term in the equations of motion:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \xi\frac{d\mathbf{r}_i}{dt}, \tag{3.11}$$

where ξ is the dynamic friction parameter connected to the deviation from the desired temperature. The equation of motion for this parameter is defined as

$$\frac{d\xi}{dt} = \frac{1}{Q}(T - T_0). \tag{3.12}$$

The strength of the coupling is determined by the parameter Q and the reference temperature T_0.

In the same spirit also the pressure coupling of the system can be described. Pressure coupling can be achieved either through the Berendsen algorithm which scales the coordinates and box vectors at every time step or through the Parrinello-Rahman approach representing the equations of motion also for the box vectors [105]. Therefore the Parrinello-Rahman pressure coupling is similar with the Nosé-Hoover scheme. The Berendsen pressure coupling relaxes the pressure in the following way:

$$\frac{d\mathbf{P}}{dt} = \frac{\mathbf{P}_0 - \mathbf{P}}{\tau_p} \tag{3.13}$$

with a scaling matrix μ given by

$$\mu_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau_p}\beta_{ij}\{P_{0ij} - P_{ij}(t)\}. \tag{3.14}$$

Here δ_{ij} is the Kronecker delta being one when i and j are equal and zero otherwise, β_{ij} is the isothermal compressibility of the system, τ_p being the time constant for pressure coupling and P_{0ij} being the target pressure. In cases where it is important to calculate the thermodynamical properties accurately, the Parrinello-Rahman pressure coupling is used. This approach is applied in this study. In the Parrinello-Rahman algorithm, the box vectors represented by the matrix $\mathbf{b}$ obey the following equation of motion:

$$\frac{d\mathbf{b}^2}{dt^2} = \frac{V}{\mathbf{W}\mathbf{b}}(\mathbf{P} - \mathbf{P}_{ref}) \tag{3.15}$$

where V is the volume of the box, $\mathbf{W}$ is a parameter which determines the strength of the coupling and $\mathbf{P}$ and $\mathbf{P}_{ref}$ represent the current and reference pressures, respectively.

3.8 Long-range interactions

In MD simulations it is important to treat the long-range interactions accurately. Long-range interactions are potentials that decay slowly, such as the electrostatic potentials. Ewald summation is a technique to efficiently sum the long-range interactions between particles and all their infinite periodic images [106]. However, this method is not suitable for large systems since the computational cost of the reciprocal term of the sum increases too much. Therefore the Particle-mesh Ewald

(PME) method was developed to improve the calculation of the reciprocal sum and in this way to increase the computational efficiency [106].

3.8.1 Ewald summation

The total electrostatic energy of N particles and the periodic images are given by

$$V = \frac{f}{2} \sum_{n_x} \sum_{n_y} \sum_{n_z} \sum_{i}^{N} \sum_{j}^{N} \frac{q_i q_j}{\mathbf{r}_{ij,\mathbf{n}}}, \tag{3.16}$$

where n_x, n_y and n_z are the box vectors and $\mathbf{r}_{ij,\mathbf{n}}$ is the real distance between the charges and not the minimun-image. The sum is very slow but conditionally convergent. In order to increase the computational efficiency, the Ewald summation method was introduced. In this method the idea is to convert the single slowly-converging sum presented in equation 3.16 into two quickly-converging terms and a constant term as follows:

$$V = V_{dir} + V_{rec} + V_0 \tag{3.17}$$

where V_{dir} represents the direct (real) sum, V_{rec} is the reciprocal (imaginary, or Fourier) sum and V_0 is the constant term. These terms are written as

$$V_{dir} = \frac{f}{2} \sum_{i,j}^{N} \sum_{n_x} \sum_{n_y} \sum_{n_z} q_i q_j \frac{\mathrm{erfc}(\beta r_{ij,\mathbf{n}})}{r_{ij,\mathbf{n}}}, \tag{3.18}$$

$$V_{rec} = \frac{f}{2\pi V} \sum_{i,j}^{N} q_i q_j \sum_{m_x} \sum_{m_y} \sum_{m_z} \frac{\exp(-(\frac{\pi \mathbf{m}}{\beta})^2 + 2\pi i \mathbf{m}(\mathbf{r}_i - \mathbf{r}_j))}{\mathbf{m}^2}, \tag{3.19}$$

$$V_0 = -\frac{f\beta}{\sqrt{\pi}} \sum_{i}^{N} q_i^2. \tag{3.20}$$

Here β determines the relative weight of the direct and reciprocal sums, V is the volume of the simulation box, $\mathbf{m}$ is a reciprocal-space vector and erfc is the error function that decreases monotonically as t increases. It is written as follows [106]:

$$\mathrm{erfc} = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} \exp(-t^2) dt \tag{3.21}$$

With these equations it is possible to use a short cutoff in both the direct and

reciprocal space sum. However, it is not realistic to use this for large systems since the computational cost of the reciprocal term increases as N^2.

3.8.2 Particle mesh Ewald-method

PME is a method developed to improve the performance of the reciprocal sum by splitting the total electrostatic energy into local interactions. The idea is to put a fine mesh on top of the system and then distribute the charges onto the mesh points [107]. The electrostatic potential is calculated by solving the Poisson equation on the mesh using the fast Fourier transformation for the reciprocal sum allowing an efficient calculation [106]. The forces on each atom are then calculated by numerically differentiating the potential and interpolating back from the mesh to the position of the atom, thus reducing memory requirements [107]. The charge interpolation function originally used was Lagrange interpolation but an enhanced PME uses the B-spline interpolation function due to the fact that it is smoother and allows higher accuracy by increasing the order of interpolation. See reference [106] for further details.

The direct sum is evaluated explicitly using cutoffs [106]. When computing the direct sum, the relative weight factor β presented in equation 3.18 is chosen large enough in order to apply a fixed cutoff radius [106]. This reduces the complexity of the direct sum. PME has several advantages in addition to the improvement in the computational cost. It is more accurate than the Ewald summation and, moreover, the accuracy can be improved by adjusting a few parameters [106]. Other ways to treat the long-range interactions include the reaction field method and the particle-particle particle-mesh method but these are not discussed further since the PME method is the one applied in this study.

3.9 Constraints

In classical MD simulations the time step is limited by bond oscillations having a relatively high frequency and low amplitude [108]. In order to increase the time step in simulations, the bonds and bond angles can be treated as constraints on the equations of motion. Constraints are algorithms implemented after the force calculations to freeze the unwanted highest frequency vibrations since the movements of covalent bonds in the system can be adjusted using constraints [98]. In other words, constraints force the bonds to stay at the proper length or angle. In addition, a constrained bond resembles more closely a quantum oscillator in its ground state than a classical oscillator, making it physically a more realistic model for a covalent bond. Examples of constraint algorithms include LINCS, SHAKE, SETTLE and RATTLE.

LINCS is a constraint algorithm that resets bonds to their correct lengths after an unconstrained update [108]. This method is non-iterative since it always uses two steps and it can therefore be easily parallelized, thus improving the computational efficiency. In addition, LINCS is more stable and faster in comparison with SHAKE but the disadvantage of this method is that it can only be used with bond constraints and isolated angle constraints. SHAKE-algorithm was introduced on the basis of the leap-frog Verlet scheme. It changes a set of unconstrained coordinates $\mathbf{r}'$ to a set of coordinates $\mathbf{r}''$ which fulfill the list of distance constraints using a set $\mathbf{r}$ as a reference [98]. Therefore it is an iterative method and the most widely used for large molecules. Compared with SHAKE, SETTLE operates in a similar manner but solves the problem analytically [108]. It avoids the linearization of the equations but requires the evaluation of ten square-root functions per tri-atomic molecule [109]. Due to this, the SETTLE algorithm is too complicated for large molecules but faster for small molecules when compared with SHAKE. RATTLE is a velocity version of the SHAKE algorithm. In addition to the calculation of the positions of the particles, RATTLE calculates the velocities at the next time step from the velocities at the present time step [110]. At each time step the algorithm guarantees that the coordinates and velocities of the atoms satisfy the internal constraints. When compared with SHAKE, RATTLE is more accurate and easier to modify for the use of MD methods with constant temperature and constant pressure since it deals directly with velocities.

4. SIMULATION MODELS AND ANALYSIS METHODS

This chapter consists of two sections. The first section describes the force fields used for different molecule types, the initial structures of the simulated systems in detail as well as the simulation parameters used. The second section lists and describes all the methods applied to analyze the obtained data from MD simulations.

4.1 Simulation systems and details

The MD simulations performed in this study serve as pioneer simulations in the field of research concerning the interactions between CETP and the novel molecular agent anacetrapib. The results of these simulations can be utilized when studying the corresponding interactions between CETP and the former developed agent torcetrapib and furthermore, when new potential molecular agents are developed. The computing facilities needed in this study were offered by the Finnish IT Centre for Scientific Computing. The simulations and their analysis were performed with the GROMACS software package, version 4.5.4 [97]. GROMACS is a widely applied simulation package for MD simulations including a large set of ready-made analysis tools.

4.1.1 Force fields and initial configurations

The coordinate file for CETP was acquired from the Protein Databank with an accession code 2OBD [57]. In addition to CETP, this file provides information of the atomic positions of the lipids involved in CETP, namely the two CEs located inside the hydrophobic tunnel and the two DOPC lipids that block the openings of the tunnel. The coordinate file of anacetrapib was obtained from Artturi Koivuniemi. The force fields for these lipids as well as for anacetrapib were obtained from Tómasz Rog whereas the force field used for CETP was the official distribution force field denoted as optimized potential for liquid simulations all-atom (OPLS-AA) [111]. The all-atom force field, as the name implies, is an explicit representation of all atoms in the simulated system, including the nonpolar hydrogens. Water molecules were described with the TIP3P model since it is compatible with the parametrization of OPLS-AA.

The starting structures of the simulations were determined and validated with the help of molecular docking calculations. The purpose was to explore the most probable sites from the crystal structure of CETP where anacetrapib desires to attach. In addition, also the optimal docking conformations of the drug can be obtained from the calculations. The program applied for these purposes was AutoDock Vina. Its technical details can be found in [112]. The program is able to select the protein-ligand complexes in terms of the lowest binding energy without prior knowledge of the binding site. In order to reduce the impact of protein flexibility, the structure of CETP was held rigid. The obtained results are presented in Figure 4.1 with ligands colored with red, brown, cyan and green representing the most probable binding sites and conformations of anacetrapib. The binding energies of these sites are -11.4 kcal mol^{-1}, -11.1 kcal mol^{-1}, -11.6 kcal mol^{-1} and -11.2 kcal mol^{-1}, respectively. The binding site of the ligand colored with brown is positioned in the hydrophobic pocket in the N-terminal domain and is mainly formed by the residues Thr23, Ile27, Ala30, Ile82, Ala191, Val194, Gln195, Arg197, Ala198 and Leu463. Some of these residues surround the binding site of the red ligand located at the N-terminal tunnel opening. Other important residues for this domain are Phe467-Leu471. The binding site of the green ligand is also located in the hydrophobic pocket but in the C-terminal domain. The surrounding residues are Ala273, Leu283, Val342, Phe344, Tyr371, Lysh373, Leu421, Phe425, Asp448 and Asp456. The cyan ligand is located on a region between the above mentioned residues. Dong et al. [113] have obtained similar results with *N,N*-disubstituted -trifluoro-3-amino-2-propanols-based inhibitors of CETP.

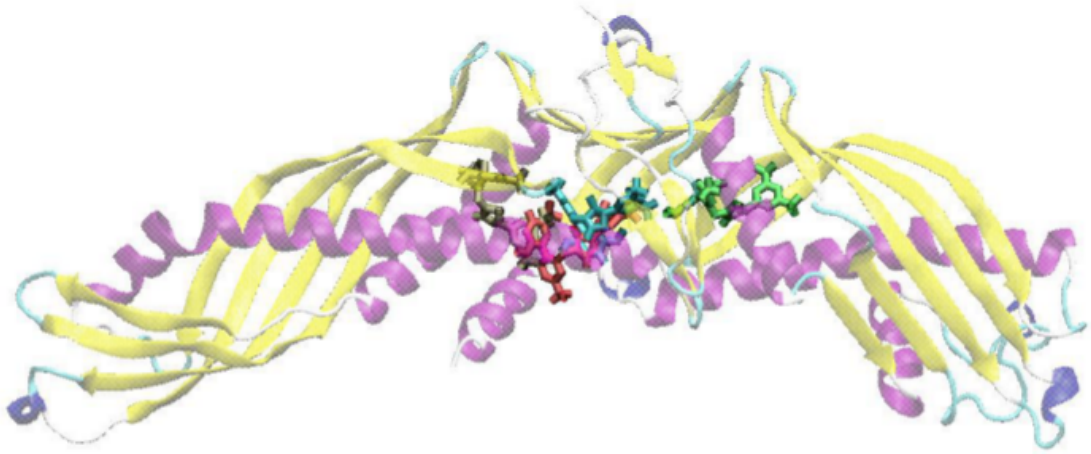


Figure 4.1: The binding sites and conformations of anacetrapib within the structure of CETP obtained from the molecular docking calculations. The binding energies for the binding sites of red, brown, cyan and green ligands are -11.4 kcal mol^{-1}, -11.1 kcal mol^{-1}, -11.6 kcal mol^{-1} and -11.2 kcal mol^{-1}, respectively.

The initial structures of the simulations were determined based on the results obtained from the molecular docking calculations presented in Figure 4.1 and they can be divided into two groups. The first group consists of four systems where anacetrapib was placed in the lipid binding pocket, whereas the second group consists of six systems including the drug placed outside of CETP but in the vicinity of the lipid binding pocket. Depending on the system, the lipid binding pocket included CEs and in addition, the openings of the tunnel were either unblocked or blocked by DOPCs. The two groups differ from each other also when concerning the simulation time. The group in which anacetrapib interacted with CETP from the lipid binding pocket was simulated for 200 ns, whereas the group with interactions occurring between the drug and the concave surface of CETP was simulated for 20 ns.

Building the initial configuration of each system required a manual placing of the molecules in order to achieve the wanted starting structure. This was done by using the visual molecular dynamics (VMD) program often used for displaying and animating biomolecular systems [114]. The construction was started from the group of longer simulated systems. The first system (L1-200ns) consists of CETP and two DOPC lipids that block the openings of the empty lipid binding pocket. The starting structure for this system was obtained directly from the Protein Databank with the removal of CEs. The starting configuration for the second system (L2-200ns) is presented in Figure 4.2. This system includes CETP, two DOPCs (red space fills) to block the openings of the tunnel as well as two anacetrapibs (cyan-white space fills) placed in the vicinity of cysteine amino acids (blue space fills) of CETP. The sites of the drug molecules correspond to the binding sites of cyan and green ligands presented in Figure 4.1. Anacetrapibs were placed in the lipid binding pocket in a way that their oxygen groups were oriented towards the cysteines since it is expected that these functional groups will form hydrogen bonds with cysteines residues thus stabilizing the CETP-anacetrapib complex. The starting configuration of the third 200 ns simulation (L3-200ns) differs from the previous one in a way that two CEs are placed in the hydrophobic tunnel of CETP instead of anacetrapibs. This structure is presented in Figure 2.5. In the fourth system (L4-200ns) DOPCs and CEs were removed from the structure and one anacetrapib was placed in the lipid binding pocket between the binding sites determined by the cyan and green ligands. Each system was solvated with approximately 108000 water molecules and neutralized with 15 sodium atoms since CETP has a net charge of -15e. The reason for neutralizing the systems is that life does not exist at net charge. In addition, the PME method described in Section 3.8 would produce artifacts if the total charge of the system differs from zero. Therefore the systems consisted of a total of 332076, 332175, 332414 and 331794 atoms, respectively.

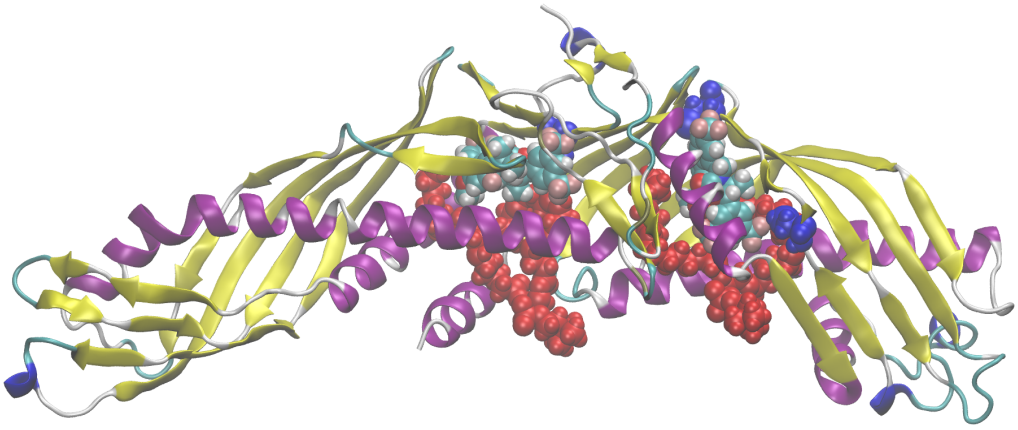


Figure 4.2: The starting structure of the second 200 ns simulation. The system includes CETP, two DOPC lipids to cover the openings of the tunnel and two anacetrapibs placed in the vicinity of cysteine amino acids. DOPCs are drawn as red space fills, anacetrapibs as cyan-white space fills and cysteine residues as blue space fills. For the sake of clarity, water is not shown.

After building the initial configurations for the longer simulations, the construction was started for the group of shorter simulations. In these simulations one anacetrapib molecule was placed outside of empty CETP. In the first simulation (S1-helix) anacetrapib was placed 1 nm distance from helix X. Concerning the four following systems (S2-1nm, S3-2nm, S4-3nm and S5-4nm), anacetrapib was placed 1, 2, 3 and 4 nm distance from the two tunnel openings, respectively. In Figure 4.3 is presented the starting configuration for the simulation S3-2nm. The sixth system (S6-convex) differs from the previous ones since anacetrapib was placed 3 nm distance from the convex back of the protein. After the systems were built, they were solvated with approximately 108000-163000 water molecules depending on the size of the simulation box. The simulation boxes in simulations S4-3nm, S5-4nm and S6-convex were larger since the distance between the particles was larger compared with simulations S1-helix, S2-1nm and S3-2nm. Hence a larger number of water molecules was needed to fill the box. The systems were neutralized in a similar manner than in the longer simulations, and thus the systems consisted of a total of 331725, 331743, 331743, 497097, 497085 and 497151 atoms, respectively. After the systems were constructed, solvated and neutralized but before any simulation was started, energy minimization was performed for each system using the steepest descent method and 500 steps. All the simulations, their abbreviations and simulation times are listed in table 4.1.

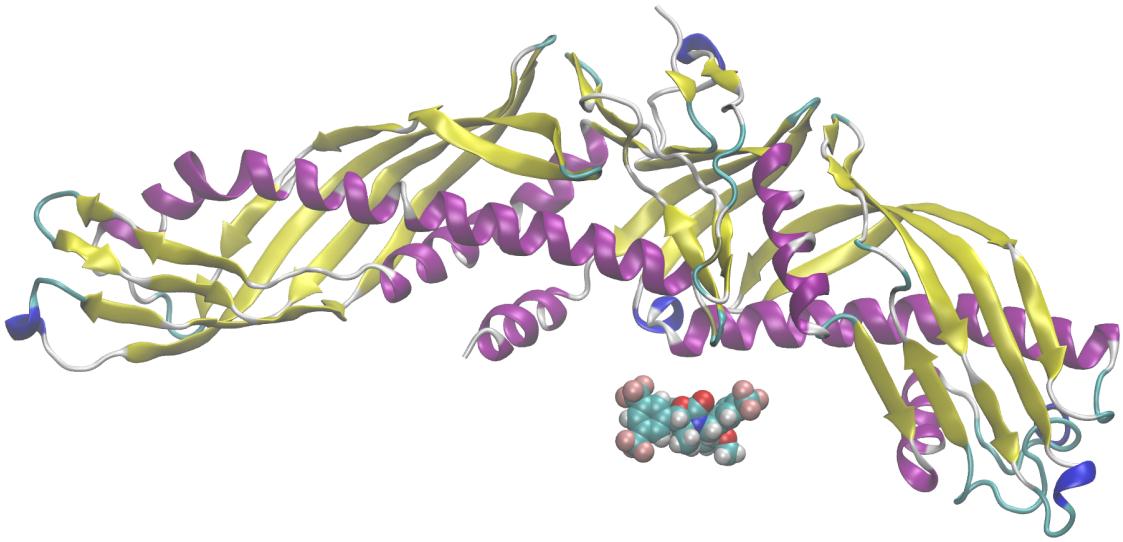


Figure 4.3: The starting structure of the third 20 ns simulation. The system includes CETP and anacetrapib placed 2 nm distance from the tunnel openings. Anacetrapib is drawn as cyan-white space fills. For the sake of clarity, water is not shown.

Concerning the simulations performed for 200 ns, the first and the third simulation (L1-200ns and L3-200ns) without anacetrapibs serve as the control simulations for the other two 200 ns simulations. The purpose of control simulations is to help the analysis of the results including anacetrapib (L2-200ns and L4-200ns), since with their help, it is possible to more easily distinguish the changes in the secondary structure of CETP induced by the drug. The shorter 20 ns simulations were constructed to reflect 'random' initial conditions where anacetrapib appears in a randomly chosen location around CETP. The reason to perform several simulations of this kind was to see the diffusion process, that is, the process of self-assembly of how anacetrapib forms a complex with CETP and whether there is a long-range force that drives the drug in the vicinity of CETP. With the help of these simulations it is also possible to evaluate the consistency of the obtained results in a manner that whether or not they end up in similar locations and final conformations defined by the red ligand in Figure 4.1. In addition, the purpose was also to evaluate the interactions between anacetrapib and the concave surface of CETP and to see the effects of anacetrapib on the conformation of helix X.

Table 4.1: The performed MD simulations, their abbreviations and simulation times.

Abbreviation	System	Time
L1-200ns	CETP, 2 x DOPC	200 ns
L2-200ns	CETP, 2 x DOPC, 2 x anacetrapib	200 ns
L3-200ns	CETP, 2 x DOPC, 2 x CE	200 ns
L4-200ns	CETP, anacetrapib	200 ns
S1-helix	CETP, anacetrapib 1 nm from helix X	20 ns
S2-1nm	CETP, anacetrapib 1 nm from tunnel openings	20 ns
S3-2nm	CETP, anacetrapib 2 nm from tunnel openings	20 ns
S4-3nm	CETP, anacetrapib 3 nm from tunnel openings	20 ns
S5-4nm	CETP, anacetrapib 4 nm from tunnel openings	20 ns
S6-convex	CETP, anacetrapib 3 nm from the convex back	20 ns

4.1.2 Simulation parameters

The actual simulations were performed under NpT conditions (constant number of particles, pressure and temperature). The reference temperature for all simulated systems was 310 K and each component of the system was separately coupled to the temperature bath using the Nosé-Hoover coupling method with time constant of 0.1 ps. The used pressure coupling method was Parrinello-Rahman barostat with isotropic pressure coupling and a coupling constant of 1 ps and the reference pressure of 1 bar. The time step of 2 fs was used for time integration and all bonds were constrained using the LINCS algorithm. The van der Waals interactions were chosen to have a cutoff at 1 nm. The long-range electrostatic interactions were evaluated by the PME method with a 1 nm cutoff for the Coulombic interactions. After completion of each simulation the respective trajectory files were analyzed with different tools of GROMACS.

4.2 Analysis methods

The structure, properties and interactions of CETP and anacetrapib can be analyzed using various methods. The observables found in MD simulations, the velocities and positions of the particles for example, can be used to express physical quantities that can be compared with experimental data, if available. This section covers the methods used to analyze the simulations and is based mostly on the manual of GROMACS [97]. All the applied analysis tools are ready-made and distributed with GROMACS.

4.2.1 Analysis of secondary structure elements

The secondary structure elements of a protein can be calculated with a define secondary structure of proteins (DSSP) program. The purpose of this analysis is to measure the most likely secondary structure of the protein as a function of time. This is achieved by reading the positions of the atoms followed by a calculation of hydrogen bond-energies between all atoms [115]. The best two hydrogen bonds for each atom are then used in order to determine the most likely class of secondary structure for each residue [115]. In other words, the DSSP program breaks down the different secondary structure elements (i.e. helix, sheet, turn, bend, etc.) of individual residues for each time step and allows the visualization of the data [116]. From this plot it is easy to see the stability (or de-stability) of secondary structure elements as a function of time.

The algorithm for extracting the different structural elements from the atomic coordinates is a process of pattern-recognition mainly based on hydrogen bonding patterns. This is due to the fact that the presence or absence of a hydrogen bond can be characterized by a single decision parameter, namely a cutoff in the bond energy. The electrostatic interaction energy between two hydrogen bonding groups can be calculated by placing partial charges on the C,O $(+q_1, -q_1)$ and N,H $(-q_2, +q_2)$ atoms written as follows [117]:

$$E = q_1 q_2 \left(\frac{1}{r_{ON}} + \frac{1}{r_{CH}} - \frac{1}{r_{OH}} - \frac{1}{r_{CN}} \right) f \tag{4.1}$$

where q_1=0.42e and q_2=0.20e, e being the unit electron charge and r_{AB} the interatomic distance between atoms A and B. f in a dimensional factor with a value of 332. The unit binding of energy E is expressed as kcal mol^{-1}. A good hydrogen bond has a binding energy of about -3 kcal mol^{-1} [117].

The patterns of hydrogen bonding define a hierarchy: n-turns with a hydrogen bond between the CO-group of residue i and the NH-group of residue $i + n$, where $n = 3, 4, 5$ and bridges with a hydrogen bond between residues that are not near each other in sequence [117]. These two types of patterns essentially cover all the backbone-backbone hydrogen bonds. α-helices and β-sheets are defined by repeating 4-turns and bridges, respectively. The different secondary structure elements are discussed in more detail in Section 3.2.

4.2.2 Root mean square deviation

The root mean square deviation (RMSD) is used to evaluate the deviation of the structure of the simulated system from the initial starting structure over the course of simulation. RMSD can be calculated by least-square fitting the simulated structure to the reference structure ($t_2 = 0$):

$$RMSD(t_1, t_2) = \sqrt{\frac{1}{M}\sum_{i=1}^{N} ||m_i \mathbf{r}_i(t_1) - \mathbf{r}_i(t_2)||^2} \tag{4.2}$$

where M is the sum of particle masses and $\mathbf{r}_i(t)$ is the position of atom i at time t. The smaller the obtained RMSD value, the smaller the deviation between the final and initial structures and moreover, the more similar the structures are. In order to determine the structural similarity, an arbitrary RMSD cutoff has been chosen to indicate that changes on the order of 0.1-0.3 nm are acceptable and expected [118].

4.2.3 Radius of gyration

Radius of gyration is used to measure the compactness of the structure of the protein as a function of time. It is mathematically defined as the mean square distance of each particle in the structure with respect to its center of mass r_{cm}. Hence it describes how the mass is divided around r_{cm}. The equation is written as

$$R_g = \sqrt{\left\langle \frac{\sum_i m_i(\mathbf{r}_i - \mathbf{r}_{cm})^2}{\sum_i m_i} \right\rangle}, \tag{4.3}$$

where the center of mass is defined as

$$\mathbf{r}_{cm} = \frac{\sum_{j=1}^{N} m_j \mathbf{r}_j}{\sum_{j=1}^{N} m_j}. \tag{4.4}$$

In this study the radius of gyration is calculated for CETP in order to measure the changes in the structure of the protein as a function of time.

4.2.4 Root mean square fluctuation

The root mean square fluctuations (RMSF) of atomic positions are used to discover and evaluate the most flexible regions of CETP. RMSFs can be calculated by fitting the simulated structure to a reference structure. The RMSF profile reveals the residues

of CETP that fluctuate the most during the simulation. Therefore, the higher the obtained RMSF value, the more the residues in question fluctuate. Typically, the most fluctuating residues of a protein can be found in loop-regions. RMSF values can also be converted into B-factor values which can then be visualized.

4.2.5 Principal component analysis

Principal component analysis (PCA), also known as the covariance analysis, is used to find correlated motions. During a simulation, both collective motions and local fluctuations occur simultaneously which complicates the distinction between the two types of motion from each other [119]. PCA helps in such cases since it can filter the global, collective (slow) motions from the local (fast) motions. It uses the covariance matrix C of the atomic coordinates:

$$C_{ij} = \left\langle M_{ii}^{\frac{1}{2}}(x_i - \langle x_i \rangle) M_{jj}^{\frac{1}{2}}(x_j - \langle x_j \rangle) \right\rangle \tag{4.5}$$

where M is a diagonal matrix containing the atom masses (mass-weighted analysis) or the unit matrix (non-mass weighted analysis) and x_i is the position of atom i. C is a symmetric 3N∗3N matrix which can be diagonalized with an orthonormal transformation matrix R defining a transformation to a new coordinate system:

$$R^T C R = diag(\lambda_1, \lambda_2, ..., \lambda_{3N}) \tag{4.6}$$

where the eigenvalue $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{3N}$. Diagonalization produces a set of eigenvectors (columns of R) and eigenvalues that describe the collective modes of fluctuations [119], where the eigenvectors represent the direction of motion and the eigenvalues represent the amount of motion along the eigenvectors [120]. Those eigenvectors that correspond to the largest eigenvalues are denoted as principal components since they represent the collective motions with largest amplitude. After diagonalization the trajectory can be projected on the eigenvectors in order to give the principal components $p_i(t)$ presented as follows:

$$\mathbf{p}(t) = R^T M^{\frac{1}{2}}(\mathbf{x}(t) - \langle \mathbf{x} \rangle). \tag{4.7}$$

The eigenvalue λ_i is the mean square fluctuation of principal component i. The first few principal components describe the collective, slow motions in the system. The trajectory can be filtered along one or more principal components. For one principal mode this is written as:

$$\mathbf{x}^f(t) = \langle \mathbf{x} \rangle + M^{-\frac{1}{2}} R_{*i} p_i(t). \tag{4.8}$$

When this analysis is performed on a macromolecule, usually the overall rotation and translation are removed in order to look at the internal motion only. This can be achieved by least square fitting to a reference structure.

4.2.6 Hydrogen bonds

Hydrogen bond formation has a remarkable role, e.g., in the stabilization of the secondary structure of a molecule. In this study the formation of hydrogen bonds is analyzed between CETP and anacetrapib in order to see where and how tightly anacetrapib binds. The formation of hydrogen bonds during a simulation can be analyzed between all possible donors D and acceptors A. OH and NH groups are regarded as donors and O is always an acceptor [121]. The determination if a hydrogen bond exists, is done by using a geometrical criterion based on the distance between hydrogen-acceptor r_{HB} and the angle between acceptor-donor-hydrogen triplet α_{HB}:

$$r \leq r_{HB} = 0.35nm, \tag{4.9}$$

$$\alpha \leq \alpha_{HB} = 30^\circ \tag{4.10}$$

There are several ways to calculate the hydrogen bonds between two groups of atoms including the calculation of donor-acceptor distance distribution of all hydrogen bonds and the hydrogen-donor-acceptor angle α distribution of all hydrogen bonds, but the one applied in this study is the calculation of the total number of hydrogen bonds in each time frame.

4.2.7 Interaction energies

Concerning the MD simulations performed for 20 ns, it is important to calculate the interaction energies in order to determine the type of non-bonded interaction that drives the diffusion process between CETP and anacetrapib. The non-bonded interactions contain a repulsion term, a dispersion term and a Coulomb term. The repulsion and dispersion terms are combined and described in terms of the weak van der Waals interactions. The Coulombic potential describes the electrostatic interactions. Van der Waals interaction energy can be calculated using the LJ potential presented in equation 3.4 and the electrostatic interaction energy using the Coulombic potential presented in equation 3.5.

4.2.8 Spatial density map

A spatial density map determines the 3D structure around a chosen molecule. It is determined based on the spatial distribution function (SDF) which indicates how much the prespecified residues move related to the rest of the molecule during the simulation. The SDF consists of two terms, namely, the radial distribution function and the radial angular distribution function. It can be written as

$$g(r) \equiv g(r_{ij}) = \langle g(r_{ij}, \omega_j) \rangle_{\omega_j}. \tag{4.11}$$

Hence it is a function of the positions r and orientations ω of all the molecules in the system. $\langle ... \rangle_{\omega_j}$ denotes the average over all orientations of the molecule j [122]. $g(r_{ij})$ represents the radial distribution function which is presented as

$$g(r_{ij}) = \frac{\langle \delta(|r_{ij}| - r) \rangle}{4\pi\rho r^2} \Delta r, \tag{4.12}$$

where ρ is the density of the molecules [123]. The term $g(\omega)$ in equation 4.11 is the radial angular distribution function written as

$$G(r, \theta) = \frac{1}{N\rho(r, \theta)} \langle \sum_i \sum_{j \neq i} \delta(r - r_{ij}) \delta(\theta - \theta_{ij}) \rangle, \tag{4.13}$$

where $\rho(r, \theta)$ is the number of molecules in the volume element presented as [123]

$$\rho(r, \theta) = \left(\frac{N}{V} \right) 2\pi r^2 \sin\theta \Delta r \Delta\theta. \tag{4.14}$$

In order to properly calculate a meaningful SDF, the particle of interest must be centered within the simulation box over the course of the simulated trajectory, and rotational and translational motions must be removed. This implies that the SDF for the surrounding species must be assembled based on a relatively fixed reference. The result obtained from the analysis can then be visualized. In this study, the SDF is determined for DOPC, CE and anacetrapib molecules in order to see how they move in the lipid binding pocket of CETP.

5. RESULTS AND DISCUSSION

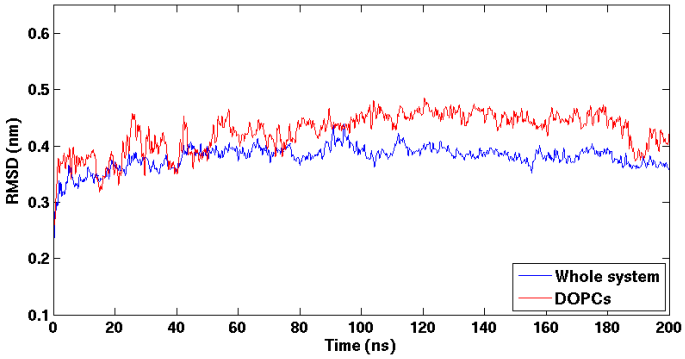
In this work, the interactions between CETP and anacetrapib were studied. A total of ten MD simulations were performed with different starting structures and simulation times. Four of the systems were simulated for 200 ns whereas six of them for 20 ns. This chapter presents the most important findings obtained from the simulations. All snapshots of simulated systems are produced using VMD and the graphs are plotted using MATLAB®.

5.1 200 ns simulations

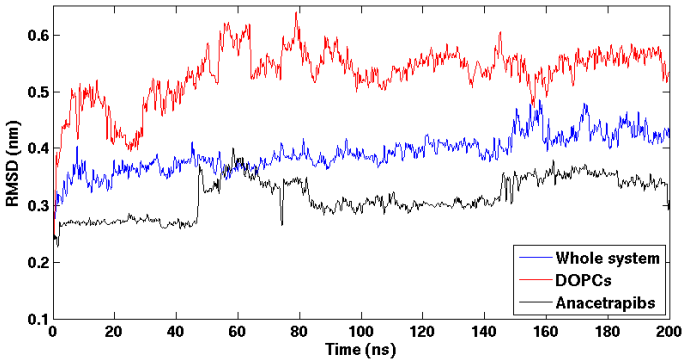
The purpose was to study the conformational changes of CETP induced by anacetrapib through the investigation of which regions of the protein were involved in the stability and which were related to structural fluctuations. In addition, regions with the biggest structural changes and their causes were examined. The changes in helix X and DOPCs were of great importance because of their crucial roles in the process of neutral lipid exchange. The simulations without anacetrapib (L1-200ns and L3-200ns) served as control simulations for the other two since the comparison between these two groups enable a more elaborate specification of the structural changes induced by the drug.

5.1.1 Generic stability

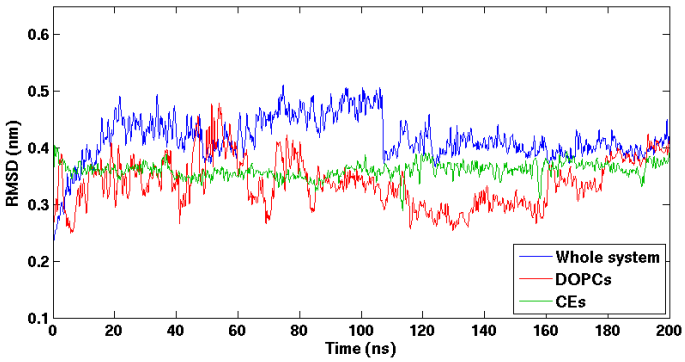
An initially random system requires time to reach a stable free energy minimum. The equilibration process of the system is affected by every external change and hence the state of equilibrium has to be observed. In this work, equilibrium is defined as a state where the large-scale motions are absent. In addition, the individual molecules are free to move around but no longer tend to move in any specific direction. The equilibration process is studied through the RMSD values plotted as a function of time in Figure 5.1. Before performing the analysis, water molecules were removed from the system in order to simplify the calculations. The graph presents the total RMSDs as well as the values for DOPCs, CEs and anacetrapibs. The profiles indicate that the structures do not deviate considerably from the X-ray structure. This is supported by data for the radii of gyration that fluctuated between 3.3 and 3.45 nm, see Figure 5.2.



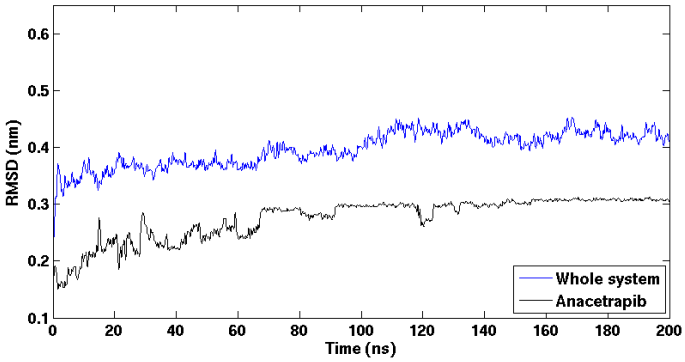
(a) L1-200ns



(b) L2-200ns



(c) L3-200ns



(d) L4-200ns

Figure 5.1: RMSD profiles of whole systems (blue) listed in table 4.1, DOPC (red), CE (green) and anacetrapib (black) molecules for the MD simulations performed for 200 ns.

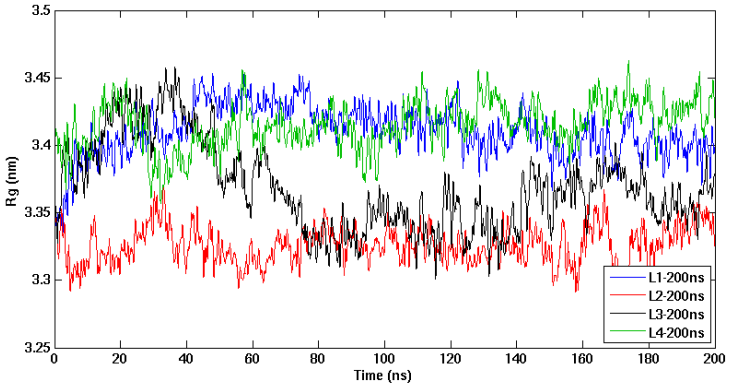


Figure 5.2: Radius of gyration profiles of MD simulations performed for 200 ns.

5.1.2 Role of phospholipids in the neutral lipid exchange

Despite the stable average structures of systems during the simulations, local fluctuations as well as structural changes occur. These alterations can be analyzed on different levels by examining individual molecules or larger molecular groups. The biggest interest is directed to DOPCs, helix X and to the changes occurring in them since these structures regulate the process of neutral lipid exchange. In addition, the influence of anacetrapib in these conformational fluctuations is of great importance due to its role in the inhibition of CETP.

The regulatory role of helix X in the lipid exchange process has been found to be significant regarding the functions of CETP [55, 59, 60]. Another crucial component in this process is PLs due to their central role both in the binding and detachment of CETP from lipoprotein surfaces. During binding, PLs have been proposed to merge into the monolayer followed by a migration away from the tunnel openings [55, 59]. This induces a formation of a hydrophobic pathway under the concave surface of the protein which permits the access of lipids into the tunnel. Concerning the detachment, the tunnel openings will need to be refilled with PLs before the dissociation since otherwise the protein would not be able to return to aqueous environment [55]. Wide attention has been paid to the mechanisms of anacetrapib affecting the regulation of these structural regions of CETP since they are regarded as essential parts of the protein's ability to change the lipids. As a consequence, the drug has been speculated to lock helix X in the open state [59, 70]. This could hinder the dissociation of the protein as well as the binding of PLs to the C-terminal opening further inhibiting CETP. However, the precise inhibitory mechanism is still unclear and thus further investigations are needed. Despite this, PLs have an evident role in the exchange and transportation of neutral lipids. This issue will be discussed in detail next.

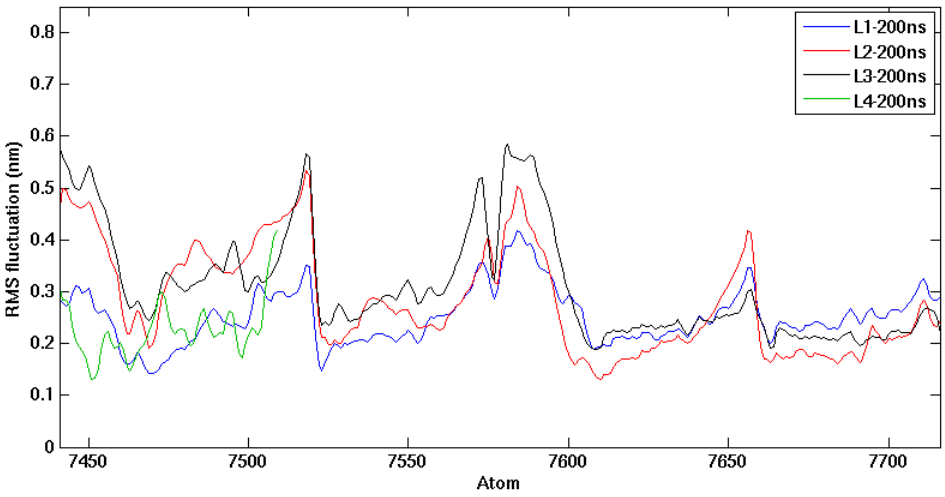
The RMSD profiles of DOPCs presented in Figure 5.1 imply increased conformational changes of DOPCs when anacetrapib is transferred into the hydrophobic tunnel. As can be seen, the RMSD fluctuates between 0.33 and 0.48 nm without the drug and between 0.4 and 0.65 nm with the drug. This result is further supported by the atomic RMSFs of DOPCs presented in Figure 5.3 as well as the corresponding spatial density maps colored with gray. The observed high fluctuations suggest that anacetrapib interacts with PLs and consequently this would destabilize the binding of DOPCs to CETP. Moreover, destabilization could reduce the ability of PLs to refill the tunnel openings stabilizing the CETP-lipoprotein complex.

Interestingly, the high fluctuations of DOPCs were observed also without anacetrapibs but when two CE molecules filled the length of the tunnel, see Figure 5.3 d. The spatial density map reveals the trajectory of DOPCs to be as wide-ranging as the one containing two anacetrapib molecules. The movement could be caused by the residence of lipids inside the hydrophobic tunnel, since in order for these molecules to properly accomodate the cavity, a conformational rearrangement of DOPCs would be required. As a consequence, some regions in the structure of PLs succumb. Either way, the movement of PLs indicates the structure of CETP to be rather unstable during the transportation of neutral lipids. This could highlight the role of helix X needed to prevent the structure of the protein from collapsing. In addition, helix X may be needed to interact with CEs in order to make their exchange possible. This is supported by several mutational studies claiming that CETP can transfer PLs but not neutral lipids without helix X [68, 69]. We will return to this issue later.

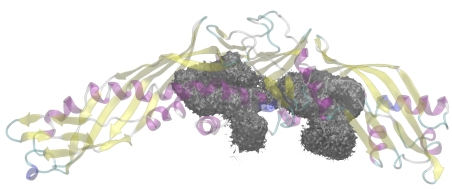
5.1.3 Mobility of helix X

RMSFs of the protein backbone were analyzed in order to find the regions that fluctuated the most during the simulations. This method of analysis can give extremely valuable information by highlighting the regions of protein backbone with low and high mobility. In addition, it provides structural information by detecting different elements of the protein secondary structure. Figure 5.4 shows the RMSFs of each simulation as a function of the residue number as well as the residual B-factors mapped to the backbone structure of CETP. As can be noticed, similar results were observed in all simulations.

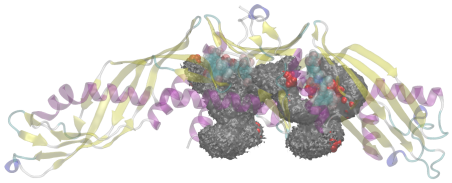
CETP has previously been reported to have mobile structures with elevated B-factors near tunnel openings including the hinge region of helix X and in the N- and C-terminal ends including the loop regions represented with omegas Ω_1 and Ω_2 [55]. As expected, these regions showed the highest mobility, with the conformational fluctuation of helix X peaking near the residue 462. In addition, five other regions in the backbone of the protein were found to fluctuate highly during each simulation. These regions were Ω_3 (residues 380-400), Ω_4 (residues 40-50), Ω_5 (residues 90-110)



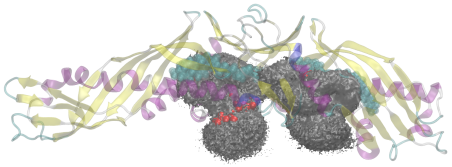
(a)



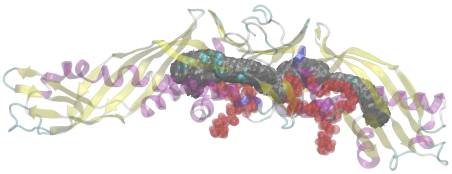
(b) L1-200ns, DOPCs



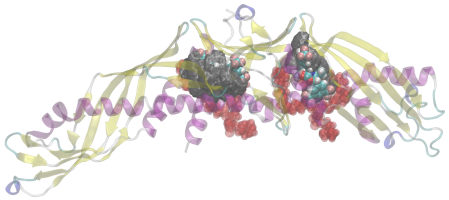
(c) L2-200ns, DOPCs



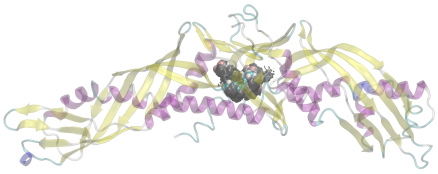
(d) L3-200ns, DOPCs



(e) L3-200ns, CEs



(f) L2-200ns, anacetrapibs



(g) L4-200ns, anacetrapib

Figure 5.3: a) RMSF profiles of MD simulations performed for 200 ns. Atoms 7441-7716 correspond to DOPCs (L1-200ns, L2-200ns and L3-200ns). The peaks indicate the atoms that fluctuated the most during the simulations. b) - g) Spatial density maps for DOPCs, CEs and anacetrapibs. The map is colored with gray revealing the movement of the molecules inside the lipid binding pocket of CETP.

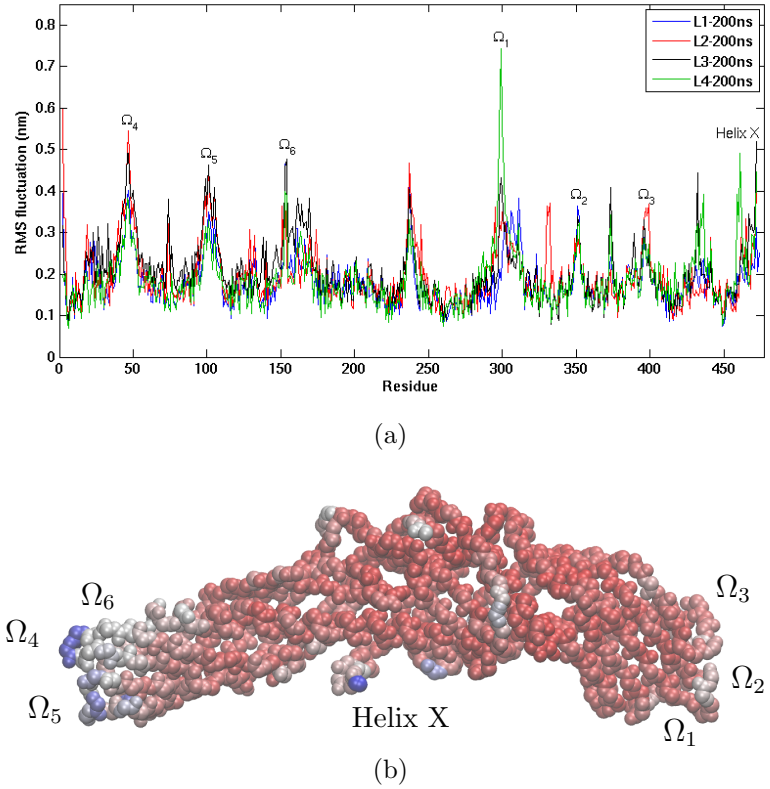


Figure 5.4: a) Residual RMS fluctuations of MD simulations performed for 200 ns. The peaks indicate the regions of CETP that fluctuated the most during the simulations. These regions are found in the loops marked with omegas as well as in the residues corresponding to helix X. b) Residual B-factors mapped to the backbone structure of CETP. Red color indicates the most rigid regions in the structure whereas white and blue indicate the most flexible structural regions. Loop regions are marked with omegas and the region corresponding to helix X is labeled.

and Ω_6 (residues 150-170) which were earlier shown to have high mobility [59]. All these regions are found in the loops and therefore the high fluctuations can be expected. The observed results were confirmed by the results obtained from PCA which showed the movement of the biggest amplitude to reside in the loops (data not shown). The results imply that the structure of CETP is elastic which could facilitate the binding to lipoprotein surfaces with varying curvatures. In addition, the observed flexibility of the hinge region could assist the lipid exchange process. Similar results were observed in [59].

5.1.4 Structure fluctuations of helix X

As mentioned, the regulatory role of helix X has been identified to play an important role in the lipid exchange process. The structure of helix X has been proposed to

undergo conformational changes during lipoprotein binding by moving aside from the N-terminal tunnel opening or by rearranging and becoming buried inside the hydrophobic pocket [59, 60]. These changes could assist the transfer of lipids between CETP and lipoprotein particles. In addition, helix X is needed to shield the tunnel opening before detachment to make CETP more compatible with the surrounding aqueous environment [55]. These observations suggest that helix X acts as a lid conducting the lipid exchange by alternating its open and closed states.

The conformational changes of helix X can be studied through the DSSP analysis which is a standard method for assigning the changes in the secondary structure to the different residues of the protein. Concerning the performed MD simulations, it became apparent that the secondary structure of helix X experienced considerable fluctuations between turn (unfolding of the helix) and 3_{10}-helix (the extension of the helix) when either CEs or anacetrapibs were present in the hydrophobic tunnel, see appendix A.1. This proposes that the drug induces conformational fluctuations to the helix which could indicate substantial effects on the lipid transfer functions of CETP.

On the basis of the above findings, the conformational rearrangement of helix X can be speculated to affect the lipid exchange in two different ways. First, the changes induced by the drug could enforce the helix to favor the open state. This would hinder the detachment of CETP from lipoprotein surfaces stabilizing the complex. In addition, it is probable that the open state may prevent CETP from binding and transporting PLs. Second, it is possible that the rearrangement locks the helix in the closed state reducing the ability of the protein to bind HDL. This suggestion is supported by the visualization of the simulation trajectory that revealed the reluctance of the helix to move aside from the tunnel opening. However, these simulations were carried out without a lipoprotein particle and hence the above discussion is only suggestive. Therefore additional simulations including the complete lipoprotein-CETP-anacetrapib complex are needed to confirm these issues.

Previous suggestions have expressed that the movement of helix X aside from the N-terminal tunnel opening is sufficient for the lipid exchange to occur [60]. However, this is in contradiction to several mutational studies that showed CETP to be capable of transfering PLs but not TGs or CEs when helix X was completely removed from the structure [68, 69]. Due to this, the role of helix X may not be as simple as previously suggested. Since helix X experienced conformational fluctuations when CEs accomodated the hydrophobic tunnel, helix X can be suggested to interact via currently unknown mechanism with CEs. This interaction could be needed for CE exchange to occur. This finding is supported by the previous MD simulation which showed that during lipoprotein binding helix X became buried inside the tunnel when two CEs were bound to CETP [59].

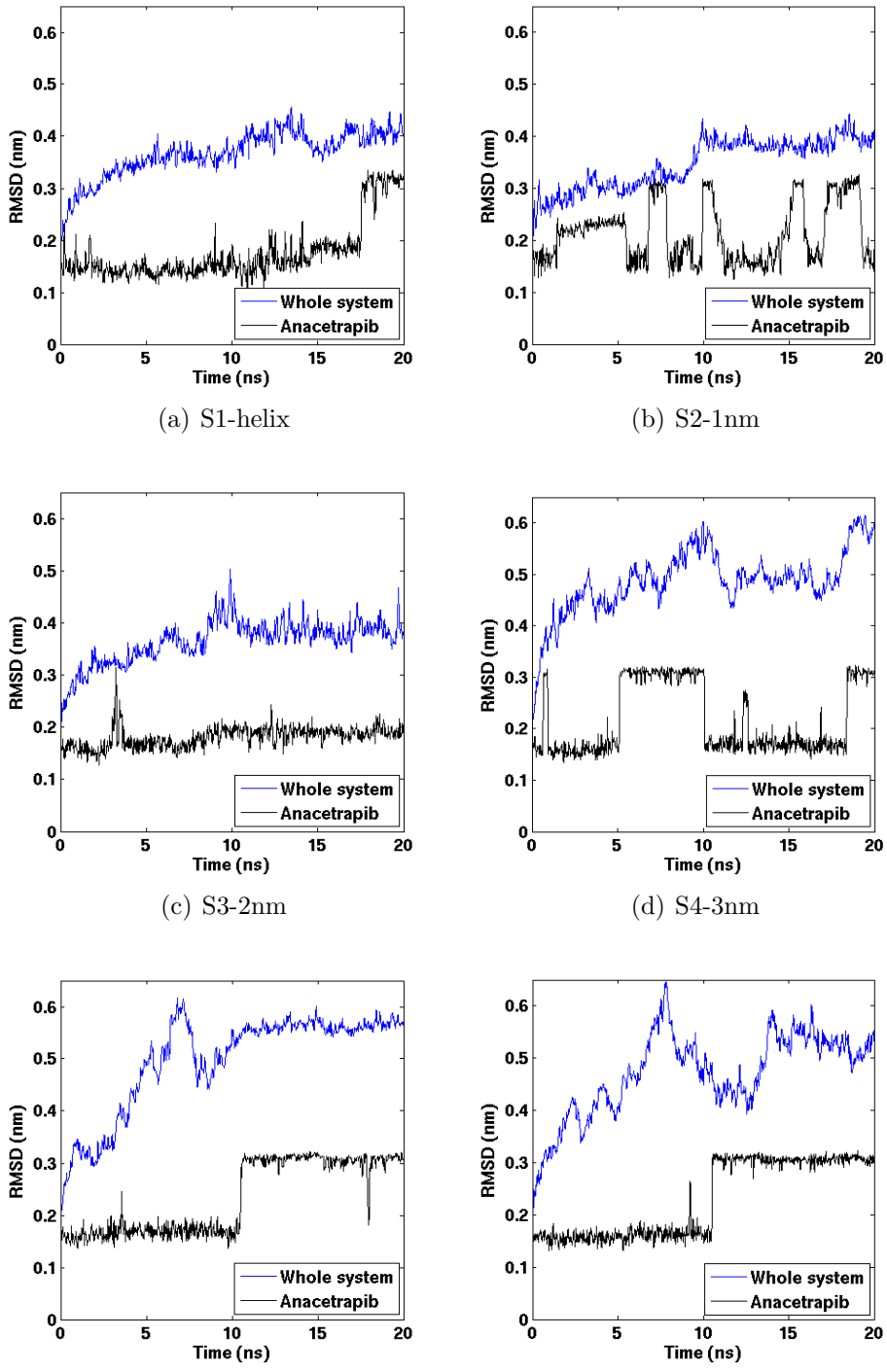
5.2 20 ns simulations

The starting structures were constructed to reflect 'random' initial conditions in order to better correspond to a biological environment. Hence anacetrapib appeared in a randomly chosen location outside but in the vicinity of CETP. The purpose was to observe the diffusion process, namely, how the two particles form a complex. In addition, the aim was to examine the existence of a long-range force that could assist the diffusion. With the aid of these simulations it would be possible to evaluate the consistency of the results. The final goal was to evaluate the interactions between anacetrapib and the concave surface of CETP and moreover, to see the possible effects of the drug to the conformation of helix X.

5.2.1 Driving force between CETP and anacetrapib

The equilibration process of each system can be observed from the RMSD profiles (without water molecules) plotted as a function of time in Figure 5.5. It is apparent that structures S1-helix, S2-1nm and S3-2nm do not deviate considerably from the corresponding X-ray structures whereas the last three systems experience fluctuations between 0.25 and 0.65 nm. The fluctuations are caused by the wandering of the drug around CETP which can be noticed from the spatial density maps presented in Figure 5.7. Wandering as well as the rotation explain the changes evident in the RMSD profiles of the drug. The radii of gyration fluctuated between 3.33 and 3.54 nm, see Figure 5.6.

The reason behind the wandering is the distance between the particles. The distance ranged from 1 to 2 nm in the first three and from 3 to 4 nm in the last three simulations. It seems that the change between distances 2 and 3 nm weakens considerably the van der Waals- (LJ potential) as well as the electrostatic (Coulombic) interactions between the particles. As can be observed from the spatial density maps presented in Figure 5.7, the movement of the drug is highly similar in S1-helix, S2-1nm and S3-2nm where anacetrapib moves at the N-terminal tunnel opening mainly around residues Arg197, Ser431, Lysh432, Gly433, Ser435 and Hisb462. These residues are in good agreement with the results obtained from molecular docking concerning especially the ligand location presented with red, see Figure 4.1. Concerning the simulations S4-3nm, S5-4nm and S6-convex, it is apparent that the strength of interactions is weak, since the drug wanders randomly around the protein. A closer inspection reveals that in S4-3nm anacetrapib ends up at 15 nm distance from CETP whereas in S5-4nm it settles in the vicinity of residues Glu155 and Gln161 located at the N-terminal end. In S6-convex anacetrapib reciprocates until it ends up in the proximity of the C-terminal end.



(a) S1-helix (b) S2-1nm

(c) S3-2nm (d) S4-3nm

(e) S5-4nm (f) S6-convex

Figure 5.5: RMSD profiles of whole systems (blue) listed in table 4.1 and anacetrapib (black) molecule for the MD simulations performed for 20 ns.

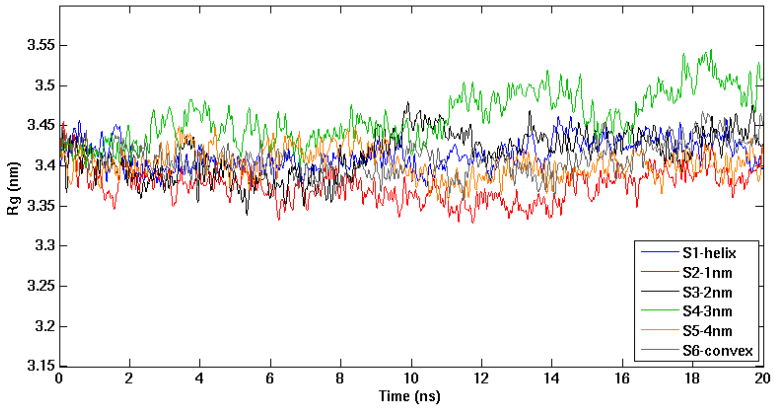
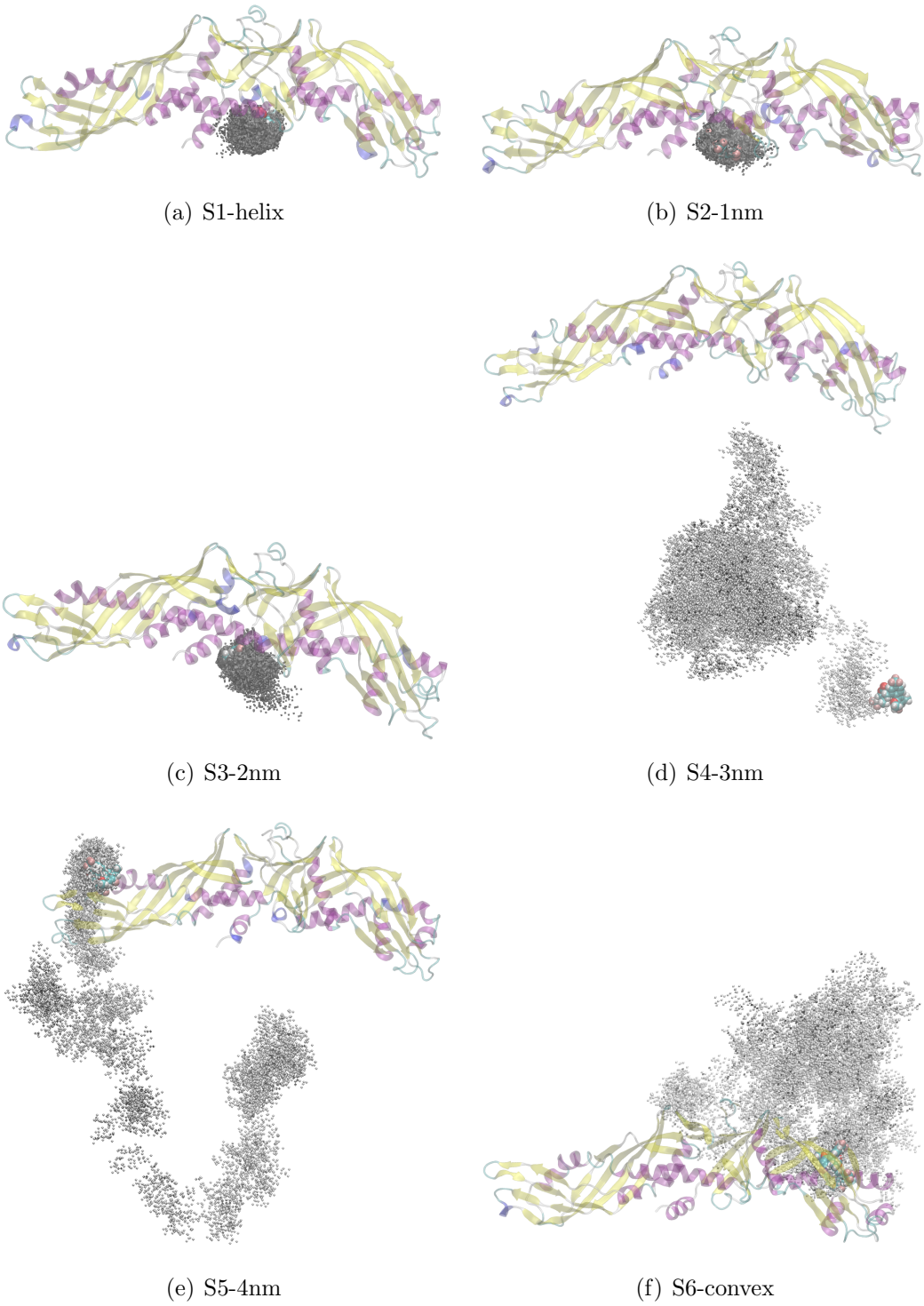


Figure 5.6: Radius of gyration profiles of MD simulations performed for 20 ns.

In addition to distance, the water phase between the particles as well as the effect of protein's charge neutralizing ions could weaken the interactions between them. These suggestions are supported by the interaction energies calculated between CETP and anacetrapib, see table 5.1 and Figure 5.8. Table 5.1 presents the interactions energies averaged over the course of simulations. It is apparent that in S1-helix, S2-1nm and S3-2nm the main force to drive the diffusion process is the weak van der Waals interaction whereas in S4-3nm, S5-4nm and S6-convex both van der Waals and electrostatic interaction forces have remarkably diminished. It seems that in S5-4nm and S6-convex anacetrapib does not interact with CETP through van der Waals interactions until 2 nm proximity. Another reason responsible for the above findings is the simulation time since 20 ns may not be long enough time to reliably interpret the results.

Table 5.1: Average interaction energies between CETP and anacetrapib. The dominant type of interaction is either weak van der Waals (E_{vdw}) interaction or electrostatic ($E_{electrostatic}$) interaction.

Simulation	E_{vdw} (kJ/mol)	$E_{electrostatic}$ (kJ/mol)
S1-helix	-105.7	-30.0
S2-1nm	-117.8	-18.4
S3-2nm	-115.8	-13.7
S4-3nm	0	0
S5-4nm	-38.4	-1.7
S6-convex	-4.5	-0.1



(a) S1-helix

(b) S2-1nm

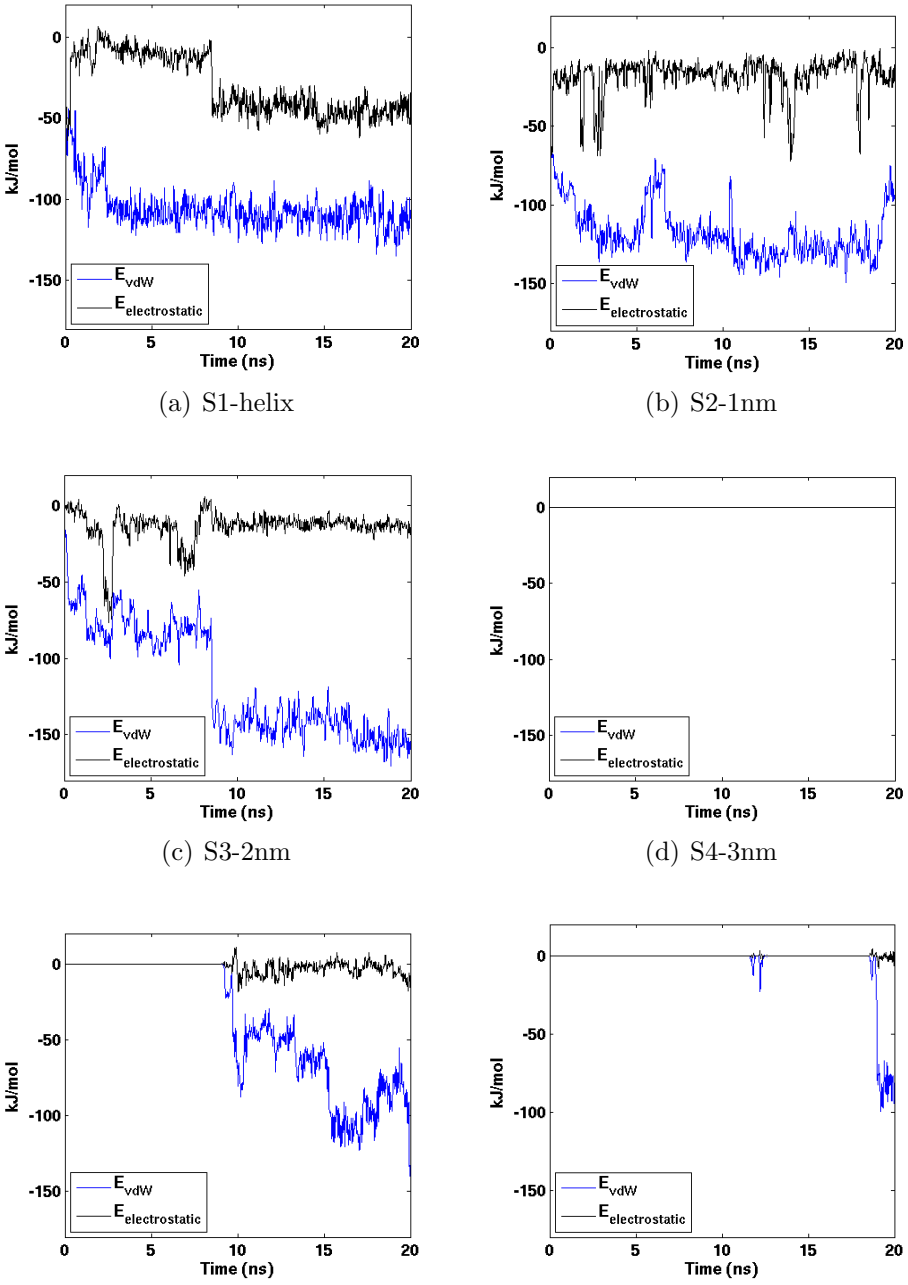
(c) S3-2nm

(d) S4-3nm

(e) S5-4nm

(f) S6-convex

Figure 5.7: Spatial density maps for anacetrapibs. The map is colored with gray revealing the movement of the drug outside the lipid binding pocket of CETP.



(a) S1-helix (b) S2-1nm

(c) S3-2nm (d) S4-3nm

(e) S5-4nm (f) S6-convex

Figure 5.8: Van der Waals and electrostatic interaction energies between CETP and anacetrapib in MD simulations performed for 20 ns.

The concept of thermal energy (k_BT) can be connected with the above mentioned interactions. Basically, thermal energy is the sum of the kinetic energies of all the atoms in a system. It originates from the individually random, disordered motion of particles. Thermal energy equals -2.6 kJ mol^{-1} in the simulation temperature (310 K) where k_B represents the Boltzmann constant. When this value is compared with the results presented in table 5.1 it is noticed that thermal energy is responsible for the movement of anacetrapib in S4-3nm, S5-4nm and S6-convex until 2 nm proximity with the protein, see Figure 5.8. As a consequence, anacetrapib wanders randomly around the protein without interacting with it.

Altogether, the affinity of anacetrapib towards the concave surface of CETP is evident taking the distance between the particles into consideration. With too large distances the movement of the drug is dominated by thermal energy resulting in disordered motion. In addition, the findings highlight the importance of electrostatic and van der Waals interactions between the drug and the protein in the formation of inhibitor-CETP complexes. The remaining question is how the attraction between the particles could be ensured in order to secure the interactions. Otherwise anacetrapib may experience random motion and may not be suitable for CETP inhibition purposes.

5.2.2 Interactions between helix X and anacetrapib

Residual RMS fluctuations for each simulation presented in Figure 5.9 indicate high mobility and increased B-factors in the hinge region of helix X as well as in the loops (marked with omegas) located at the N- and C-terminal ends. When these results are compared with the RMSFs of longer simulations shown in Figure 5.4 it is explicit that the results are similar. Additionally, equivalent regions were observed to experience fluctuations regarding the PCA method (data not shown). These results verify the previously suggested structural elasticity of CETP and its ability to bind lipoproteins with varying curvatures.

As presented, the significant conformational fluctuations of helix X were observed in the 200 ns simulations. In accordance with these findings, similar results were discerned regarding the shorter simulations since helix X was noticed to fluctuate between turn and 3_{10}-helix in S1-helix, S2-1nm and S3-2nm, see appendix A.2. Anacetrapib was noticed to move at the N-terminal opening which indicates that the secondary structure changes may be induced by the drug. Based on the above findings, it is tempting to speculate the mechanisms of the drug related to CETP inhibition. As it seems that the drug stays at the N-terminal opening, it is possible that it could prevent the transfer of lipids through the closure of the corresponding tunnel opening. Secondly, the possible staying at the tunnel opening could prevent the binding of PLs onto the surface of a lipoprotein. This could hinder the complex

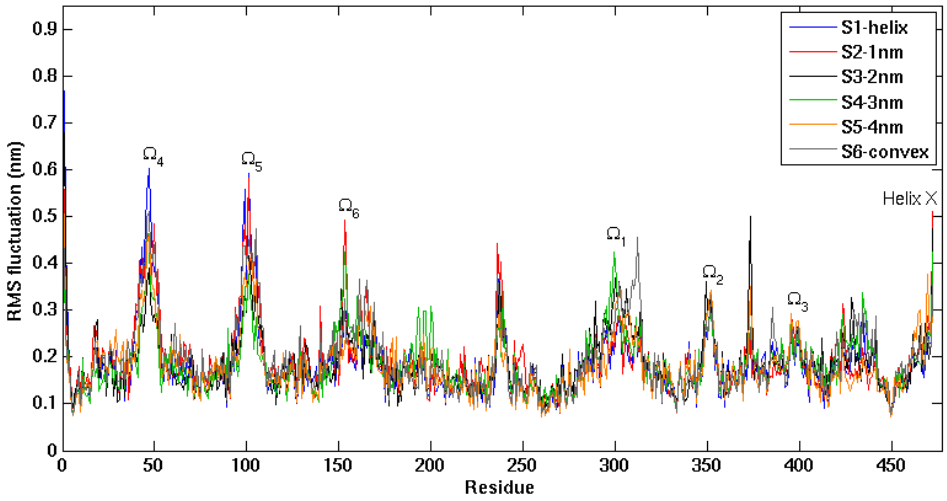
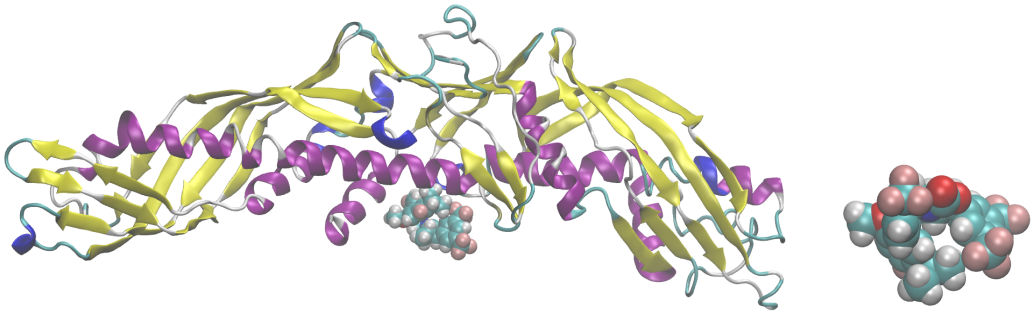


Figure 5.9: Residual RMS fluctuations of MD simulations performed for 20 ns. The peaks indicate the regions of CETP that fluctuated the most during the simulations. These regions are found in the loops marked with omegas as well as in the residues corresponding to helix X.

formation between CETP and a lipoprotein particle further reducing the ability of CETP to transfer neutral lipids.

The MD simulations performed for 200 ns suggested that anacetrapib could induce changes to the secondary structure of helix X when it is transferred to the hydrophobic tunnel of CETP. The visualization of the simulation trajectories of the 20 ns simulations revealed that when the drug moved at the N-terminal tunnel opening, the trifluoromethyl- and methylene groups of the drug oriented close to each other, see Figure 5.10. This indicates that the drug aligns itself to a tighter conformation suggesting a possible movement into the tunnel. Hence anacetrapib could prefer to interact with helix X as well as with PLs from the hydrophobic cavity. The suggested movement is supported by the lack of hydrogen bonds between the protein and the drug. Anacetrapib was noticed to form separate hydrogen bonds with the residues locating at the N-terminal opening in S1-helix, S2-1nm and S3-2nm. This implies a weak binding of the drug to CETP since when a new bond was formed, the previous one was broken. This implies an unobstructed movement of the drug. However, the binding affinity may have been affected by the simulation time which may not be long enough to fully confirm these propositions.

Despite the promising results observed in this study, several issues concerning the structure-function relationship of PLs and helix X as well as the interactions between anacetrapib, CETP and different lipoprotein particles require further investigations. Specifically the functions of helix X and PLs during lipoprotein binding and exchange of lipids are of great importance due to their essential roles in these processes.



(a) (b)

Figure 5.10: Snaphots from MD simulations performed for 20 ns. a) Snapshot of anacetrapib at the N-terminal tunnel opening. b) Anacetrapib rotated for 90° from a). The snapshot reveals the proximity of trifluoromethyl- and methylene groups from each other.

Additionally, attention should be paid to the interactions between helix X and CEs in order to explain how CETP can exchange PLs but not CEs or TGs in the absence of helix X. The novel molecular agent anacetrapib developed for CETP inhibition purposes could have a central role in all of the above mentioned interactions and hence it is crucial to continue the research related to it. Both anacetrapib and torcetrapib have previously been reported to increase the binding affinity of CETP towards lipoproteins and to inhibit the exchange of lipids [70, 73]. However, the precise mechanisms behind the increased binding affinity as well as inhibition require further clarification and therefore the research could be extended to involve a lipoprotein particle in the simulations. The ultimate goal in the future is to combine computer simulations and experimental studies to better understand the structure-function relationship of the HDL-CETP-anacetrapib complex.

6. CONCLUSIONS

The purpose of this work was to study the interactions between CETP and anacetrapib with atomistic molecular dynamics simulations. In addition, attention was paid to the structure-function relationship of PLs and helix X. A total of ten different simulation systems were constructed. They can be divided into two groups based on their starting configurations and simulation times. The first group consisted of four systems where the drug molecule was placed inside the lipid binding pocket of the protein. Each of these systems were simulated for 200 ns. In the second group the drug was placed outside but in the vicinity of the hydrophobic pocket to reflect 'random' initial conditions and hence to better correspond to the biological environment. These systems were simulated for 20 ns. Previously atomic scale simulations have been used in the studies concerning CETP and its functions but all of them have focused on the investigation of interactions between the protein and a lipoprotein particle ignoring the effects of the drug.

The focus of this work is on the conformational changes experienced by helix X and PLs during the simulations as well as on the effects of anacetrapib related to these changes. Additionally, the aim was to examine the diffusion process between the protein and the drug. Several clinical trials have shown that anacetrapib is able to inhibit the lipid transfer functions of CETP and as a consequence, to increase the level of HDL-C and to decrease the level of LDL-C [7]. The cholesterol level changes of this kind correlate with a reduced risk of cardiovascular diseases. Despite these promising results, the precise inhibitory mechanism of the drug still needs elucidation. Nevertheless, it has been proposed that the drug increases the binding affinity between CETP and different lipoprotein particles [70]. In addition, anacetrapib has been suggested either to hinder the binding or the detachment of the protein from lipoprotein surfaces thus preventing the lipid transfer [70]. This work attempts to clarify the mechanisms behind these functions as well as to characterize the effects of anacetrapib on the capability of CETP to transport and exchange neutral lipids.

The results obtained in this study point towards the important roles of helix X and PLs related to the process of neutral lipid exchange. These structures were noticed to experience considerable conformational fluctuations induced by anacetrapib. Due to this, the results emphasize the importance of the drug in regulating the functions of these structures and pave the way for the speculation about the capability of the drug

to inhibit CETP. However, it is important to notice that the simulations excluded a lipoprotein particle, and hence these results are only suggestive. Nevertheless, on the basis of the simulations, the affinity between the drug and the protein was evident when the distance between the particles is taken into consideration. In addition to distance, the results highlighted the importance of electrostatic and van der Waals interactions in the diffusion process. The remaining question is how the affinity could be ensured in order to secure the interactions between the particles. Otherwise anacetrapib may experience random motion and may not be suitable for CETP inhibition purposes. Despite the promising results observed in this study, additional simulations including the complete lipoprotein-CETP-anacetrapib complex are required to further clarify the structure-function relationships of helix X and PLs as well as the inhibitory mechanism of the drug.

This work is mostly pioneering research concerning the interactions between CETP and anacetrapib and thus paves the way to extend the scope of computational studies to gain a much deeper understanding concerning the inhibition of CETP. It is important to extend the research to involve an HDL particle in the simulations in order to better clarify the mechanisms behind the inhibitory function of the drug as well as the interactions between helix X and PLs needed for the lipid exchange to occur. The ultimate goal in the future is to combine the computer simulations and experimental studies to validate the complicated interactions between the complete HDL-CETP-anacetrapib complex. The novel understanding could be used to develop new molecular agents to fight against the progression of cardiovascular diseases since the existing medicines are not able to prevent all cardiovascular events. Overall, this work gives a solid foundation for future research projects related to anacetrapib.

REFERENCES

[1] **K. Kristiansson**. Genetics of cardiovascular disease: A candidate gene study of USF1. 8, 2008.

[2] **P. Libby**. Atherosclerosis: The new view. *Scientific American*, 286:46–55, 2002.

[3] **Suomen Sydänliitty ry**. Suositus veren kolesteroli-ja rasva-arvoista. `http://arkisto.sydanliitto.fi/kaikki_sydamesta/kolesteroli/fi_FI/suositus/`, 2005. Retrieved on July 26, 2011.

[4] **A.J. Lusis**. Atherosclerosis. *Nature*, 407:233–241, 2000.

[5] **A.R. Tall**. Plasma cholesteryl ester transfer protein. *Journal of Lipid Research*, 34:1255–1274, 1993.

[6] **P.J. Barter, M. Caulfield, M. Eriksson, S.M. Grundy, J.J.P. Kastelein, M. Komajda, J. Lopez-Sendon, L. Mosca, J.C. Tardif, D.D. Waters, C.L. Shear, J.H. Revkin, K.A. Buhr, M.R. Fisher, A.R. Tall, and B. Brewer**. Effects of torcetrapib in patients at high risk for coronary events. *New England Journal of Medicine*, 357:2109–2122, 2007.

[7] **C.P. Cannon, S. Shah, H.M. Dansky, M. Davidson, E.A. Brinton, A.M. Gotto Jr, M. Stepanavage, S.X. Liu, P. Gibbons, T.B. Ashraf, J. Zafarino, Y. Mitchel, and P. Barter**. Safety of anacetrapib in patients with or at high risk for coronary heart disease. *New England Journal of Medicine*, 363:2406–2415, 2010.

[8] Sydän- ja verisuonisairauksien ja diabeteksen asiantuntijaryhmän raportti 2008. *Kansanterveyslaitoksen julkaisuja B*, pages 1–62, 2008.

[9] **Cardiovascular disease foundation**. `http://www.cvdf.org/default.aspx`. Retrieved on July 8, 2011.

[10] **P. Barter, J. Kastelein, A. Nunn, and R. Hobbs**. High density lipoproteins (HDLs) and atherosclerosis; The unanswered questions. *Atherosclerosis*, 168:195–211, 2003.

[11] **P. Natarajan, K.K. Ray, and C.P. Cannon**. High-density lipoprotein and coronary heart disease: Current and future therapies. *Journal of the American College of Cardiology*, 55:1283–1299, 2010.

[12] **Suomen Sydänliitty ry**. Kolesterolia alentavat lääkkeet. `http://www.sydanliitto.fi/kolesterolia-alentavat-laakkeet`, 2005. Retrieved on September 6, 2011.

[13] **R.J. Chilton**. Pathophysiology of coronary heart disease: A brief review. *Journal of the American Osteopathic Association*, 104:S5–S8, 2004.

[14] **P. Libby**. Current concepts of the pathogenesis of the acute coronary syndromes. *Circulation*, 104:364–372, 2001.

[15] **K. Laitinen**. Mitkä ovat statiinien haitta- ja yhteisvaikutukset? *Apteekkari*, 8:18–19, 2004.

[16] **S. O'Sullivan**. Statins: A review of benefits and risks. *Trinity Student Medical Journal*, 8:52–56, 2007.

[17] **P. Libby, P.M. Ridker, and A. Maseri**. Inflammation and atherosclerosis. *Circulation*, 105:1135–1143, 2002.

[18] **D. Polk and P.K. Shah**. Cholesterol ester transfer protein (CETP) and atherosclerosis. *Drug Discovery Today: Therapeutic Strategies*, 4:137–145, 2007.

[19] **T.N. Tulenko and A.E. Sumner**. The physiology of lipoproteins. *Journal of nuclear cardiology*, 9:638–649, 2002.

[20] **M.H. Davidson and R.S. Rosenson**. Novel targets that affect high-density lipoprotein metabolism: The next frontier. *The American journal of cardiology*, 104:52E–57E, 2009.

[21] **G.S. Berenson, S.R. Srinivasan, B. Weihang, W.P. Newman, R.E. Tracy, and W.A. Wattigney**. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. *The New England journal of medicine*, 338:1650–1656, 1998.

[22] **Terveyskirjasto Duodecim**. Kolesteroli. `http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti?p_artikkeli=dlk00035`, 2010. Retrieved on September 6, 2011.

[23] **Patient UK Newspaper**. Preventing cardiovascular diseases. `http://www.patient.co.uk/pdf/pilsL82.pdf`, 2010. Retrieved on July 8, 2011.

[24] **J. Huttunen, A. Kesäniemi, P. Kovanen, T. Kuusi, M. Laakso, A. Lehtonen, T. Miettinen, P. Pentikäinen, K. Pyörälä, A. Reunanen, J. Salmi, M. Salo, T. Strandberg, M. Syvänne, M.R. Taskinen,**

M. Tikkanen, R. Tilvis, A. Tuomi, M. Uusitupa, and J. Viikari. *Lipidiaineenvaihdunnan häiriöt*. Orion.

[25] **T.J. Wang, M.J. Pencina, S.L. Booth, P.F. Jacques, E. Ingelsson, K. Lanier, E.J. Benjamin, R.B. D'Agostino, M. Wolf, and R.S. Vasan**. Vitamin D deficiency and risk of cardiovascular disease. *Circulation*, 117:503–511, 2008.

[26] **E.D. Michos, J.P. Reis, and M.L. Melamed**. Vitamin D status and cardiovascular health: A 2009 update. *The open clinical chemistry journal*, 3: 51–59, 2010.

[27] **C.J. Lavie, J.H. Lee, and R.V. Milani**. Vitamin D and cardiovascular disease: Will it live up to its hype? *Journal of the American College of Cardiology*, 58:1547–1556, 2011.

[28] **J.B. German, J.T. Smilowitz, and A.M. Zivkovic**. Lipoproteins: When size really matters. *Current opinion in colloid & interface science*, 11:171–183, 2006.

[29] **H.K. Walker, W.D. Hall, and J.W. Hurst**. *Clinical methods. The history, physical, and laboratory examinations*. Butterworths, Boston, 3rd edition, 1990.

[30] **T. Hevonoja, M.O. Pentikäinen, M.T Hyvönen, P.T. Kovanen, and M. Ala-Korpela**. Structure of low density lipoprotein (LDL) particles: Basis for understanding molecular changes in modified LDL. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, 1488:189–210, 2000.

[31] **T. Murtola, T.A. Vuorela, M.T. Hyvönen, S.J. Marrink, M. Karttunen, and I. Vattulainen**. Low density lipoprotein: Structure, dynamics, and interactions of apoB-100 with lipids. *Soft Matter*, 7:8135–8141, 2011.

[32] **S.M. Grundy**. Role of low-density lipoproteins in atherogenesis and development of coronary heart disease. *Clinical chemistry*, 41:139–146, 1995.

[33] **O.G. Mouritsen**. *Life-as a matter of fat*. Springer Berlin, Germany, 2005.

[34] **W.P. Castelli**. Cholesterol and lipids in the risk of coronary artery disease–the Framingham Heart Study. *The Canadian journal of cardiology*, 4:5A, 1988.

[35] **J.D. Curb, R.D. Abbott, B.L. Rodriguez, K. Masaki, R. Chen, D.S. Sharp, and A.R. Tall**. A prospective study of HDL-C and cholesteryl ester transfer protein gene mutations and the risk of coronary heart disease in the elderly. *Journal of lipid research*, 45:948–953, 2004.

[36] **R. Vikstedt**. The role of phospholipid transfer protein and cholesteryl ester transfer protein in reverse cholesterol transport. 2009.

[37] **C.G. Santos-Gallego, B. Ibanez, and J.J. Badimon**. HDL-cholesterol: Is it really good? Differences between apoA-I and HDL. *Biochemical pharmacology*, 76:443–452, 2008.

[38] **H. Saito, S. Lund-Katz, and M.C. Phillips**. Contributions of domain structure and lipid interaction to the functionality of exchangeable human apolipoproteins. *Progress in Lipid Research*, 43:350–380, 2004.

[39] **T. Vuorela, A. Catte, P.S. Niemelä, A. Hall, M.T. Hyvönen, S.J. Marrink, M. Karttunen, and I. Vattulainen**. Role of lipids in spheroidal high density lipoproteins. *PLoS Computational Biology*, 6:e1000964, 2010.

[40] **A. Catte, J.C. Patterson, D. Bashtovyy, M.K. Jones, F. Gu, L. Li, A. Rampioni, D. Sengupta, T. Vuorela, P. Niemelä, S.J. Marrink, I. Vattulainen, and J.P. Segrest**. Structure of spheroidal HDL particles revealed by combined atomistic and coarse-grained simulations. *Biophysical journal*, 94:2306–2319, 2008.

[41] **C. Bruce, R.A. Chouinard Jr, and A.R. Tall**. Plasma lipid transfer proteins, high-density lipoproteins, and reverse cholesterol transport. *Annual review of nutrition*, 18:297–330, 1998.

[42] **J.P. Segrest, M.K. Jones, A.E. Klon, C.J. Sheldahl, M. Hellinger, H. De Loof, and S.C. Harvey**. A detailed molecular belt model for apolipoprotein AI in discoidal high density lipoprotein. *Journal of Biological Chemistry*, 274:31755–31758, 1999.

[43] **A.Y. Shih, P.L. Freddolino, A. Arkhipov, and K. Schulten**. Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. *Journal of structural biology*, 157:579–592, 2007.

[44] **A.Y. Shih, S.G. Sligar, and K. Schulten**. Maturation of high-density lipoproteins. *Journal of The Royal Society Interface*, 6:863–871, 2009.

[45] **R.A. Silva, R. Huang, J. Morris, J. Fang, E.O. Gracheva, G. Ren, A. Kontush, W.G. Jerome, K.A. Rye, and W.S. Davidson**. Structure of apolipoprotein AI in spherical high density lipoproteins of different sizes. *Proceedings of the National Academy of Sciences*, 105:12176–12181, 2008.

[46] **L. Yetukuri, S. Söderlund, A. Koivuniemi, T. Seppänen-Laakso, P.S. Niemelä, M. Hyvönen, M.R. Taskinen, I. Vattulainen, M. Jauhiainen, and M. Orešič**. Composition and lipid spatial distribution of HDL particles in subjects with low and high HDL-cholesterol. *Journal of lipid research*, 51:2341–2351, 2010.

[47] **E.A. Kaperonis, C.D. Liapis, J.D. Kakisis, D. Dimitroulis, and V.G. Papavassiliou**. Inflammation and atherosclerosis. *European journal of vascular and endovascular surgery*, 31:386–393, 2006.

[48] **L. Lind**. Circulating markers of inflammation and atherosclerosis. *Atherosclerosis*, 169:203–214, 2003.

[49] **K. Ikewaki, D.J. Rader, T. Sakamoto, M. Nishiwaki, N. Wakimoto, J.R. Schaefer, T. Ishikawa, T. Fairwell, L.A. Zech, H. Nakamura, M. Nagano, and H.B. Brewer**. Delayed catabolism of high density lipoprotein apolipoproteins AI and A-II in human cholesteryl ester transfer protein deficiency. *Journal of Clinical Investigation*, 92:1650–1658, 1993.

[50] **M.L. Brown, A. Inazu, C.B. Hesler, L.B. Agellon, C. Mann, M.E. Whitlock, Y.L. Marcel, R.W. Milne, J. Koizumi, H. Mabuchi, R. Takeda, and A.R. Tall**. Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature*, 342: 448–451, 1989.

[51] **S. Zhong, D.S. Sharp, J.S. Grove, C. Bruce, K. Yano, J.D. Curb, and A.R. Tall**. Increased coronary heart disease in Japanese-American men with mutation in the cholesteryl ester transfer protein gene despite increased HDL levels. *Journal of Clinical Investigation*, 97:2917–2923, 1996.

[52] **P.J. Barter, H.B. Brewer, M.J. Chapman, C.H. Hennekens, D.J. Rader, and A.R. Tall**. Cholesteryl ester transfer protein: A novel target for raising HDL and inhibiting atherosclerosis. *Arteriosclerosis, thrombosis, and vascular biology*, 23:160–167, 2003.

[53] **M.J. Chapman, W.L. Goff, M. Guerin, and A. Kontush**. Cholesteryl ester transfer protein: At the heart of the action of lipid-modulating therapy with statins, fibrates, niacin, and cholesteryl ester transfer protein inhibitors. *European Heart Journal*, 31:149–164, 2010.

[54] **P. Roy, R. MacKenzie, T. Hirama, XC Jiang, P. Kussie, A. Tall, E. Rassart, and R. Milne**. Structure-function relationships of human

cholesteryl ester transfer protein: Analysis using monoclonal antibodies. *Journal of lipid research*, 37:22–34, 1996.

[55] **X. Qiu, A. Mistry, M.J. Ammirati, B.A. Chrunyk, R.W. Clark, Y. Cong, J.S. Culp, D.E. Danley, T.B. Freeman, K.F. Geoghegan, et al.** Crystal structure of cholesteryl ester transfer protein reveals a long tunnel and four bound lipid molecules. *Nature structural & molecular biology*, 14:106–113, 2007.

[56] **A.J. Lusis, S. Zollman, R.S. Sparkes, I. Klisak, T. Mohandas, D. Drayna, and R.M. Lawn**. Assignment of the human gene for cholesteryl ester transfer protein to chromosome 16q12-16q21. *Genomics*, 1:232–235, 1987.

[57] **Protein data bank**. Crystal structure of cholesteryl ester transfer protein. `http://www.pdb.org/pdb/results/results.do?outformat=&qrid=EDE455F1&tabtoshow=Current`. Retrieved on July 20, 2011.

[58] **Z. Yuan, T.L. Bailey, and R.D. Teasdale**. Prediction of protein B-factor profiles. *Proteins: Structure, Function, and Bioinformatics*, 58:905–912, 2005.

[59] **A. Koivuniemi, T. Vuorela, P.T. Kovanen, I. Vattulainen, and M.T. Hyvönen**. Lipid exchange mechanism of the cholesteryl ester transfer protein clarified by atomistic and coarse-grained simulations. *PLoS computational biology*, 8, 2012.

[60] **C. Bruce, L.J. Beamer, and A.R. Tall**. The implications of the structure of the bactericidal/permeability-increasing protein on the lipid-transfer function of the cholesteryl ester transfer protein. *Current opinion in structural biology*, 8:426–434, 1998.

[61] **N.M. Pattnaik and D.B. Zilversmit**. Interaction of cholesteryl ester exchange protein with human plasma lipoproteins and phospholipid vesicles. *Journal of Biological Chemistry*, 254:2782–2786, 1979.

[62] **R.B. Weinberg, V.R. Cook, J.B. Jones, P. Kussie, and A.R. Tall**. Interfacial properties of recombinant human cholesterol ester transfer protein. *Journal of Biological Chemistry*, 269:29588–29591, 1994.

[63] **R.E. Morton and D.J. Greene**. The surface cholesteryl ester content of donor and acceptor particles regulates CETP. *Journal of lipid research*, 44: 1364–1372, 2003.

[64] **B.F. Asztalos, K.V. Horvath, K. Kajinami, C. Nartsupha, C.E. Cox, M. Batista, E.J. Schaefer, A. Inazu, and H. Mabuchi**. Apolipoprotein

composition of HDL in cholesteryl ester transfer protein deficiency. *Journal of lipid research*, 45:448–455, 2004.

[65] **M.H. Davidson**. Update on CETP inhibition. *Journal of Clinical Lipidology*, 4:394–398, 2010.

[66] **R.W. Clark, R.B. Ruggeri, D. Cunningham, and M.J. Bamberger**. Description of the torcetrapib series of cholesteryl ester transfer protein inhibitors, including mechanism of action. *Journal of lipid research*, 47:537–552, 2006.

[67] **A.R. Tall**. Functions of cholesterol ester transfer protein and relationship to coronary artery disease risk. *Journal of Clinical Lipidology*, 4:389–393, 2010.

[68] **S. Wang, L. Deng, R.W. Milne, and A.R. Tall**. Identification of a sequence within the c-terminal 26 amino acids of cholesteryl ester transfer protein responsible for binding a neutralizing monoclonal antibody and necessary for neutral lipid transfer activity. *Journal of Biological Chemistry*, 267: 17487–17490, 1992.

[69] **S. Wang, X. Wang, L. Deng, E. Rassart, R.W. Milne, and A.R. Tall**. Point mutagenesis of carboxyl-terminal amino acids of cholesteryl ester transfer protein. opposite faces of an amphipathic helix important for cholesteryl ester transfer or for binding neutralizing antibody. *Journal of Biological Chemistry*, 268:1955–1959, 1993.

[70] **D. Masson**. Anacetrapib, a cholesterol ester transfer protein (CETP) inhibitor for the treatment of atherosclerosis. *Current opinion in investigational drugs (London, England: 2000)*, 10:980–987, 2009.

[71] **American Medical Association**. Statement on a nonproprietary name adopted by the USAN council: Anacetrapib. `http://www.ama-assn.org/ama1/pub/upload/mm/365/anacetrapib.pdf`, 2007. Retrieved on July 25, 2011.

[72] **C.J. Smith, A. Ali, M.L. Hammond, H. Li, Z. Lu, J. Napolitano, G.E. Taylor, C.F. Thompson, M.S. Anderson, Y. Chen, S.S. Eveland, Q. Guo, S.A. Hyland, D.P. Milot, C.P. Sparrow, S.D. Wright, A.M. Cumiskey, M. Latham, L.B. Peterson, R. Rosa, J.V. Pivnichny, X. Tong, S.S. Xu, and P.J. Sinclair**. Biphenyl-substituted oxazolidinones as cholesteryl ester transfer protein inhibitors: Modifications of the oxazolidinone ring leading to the discovery of anacetrapib. *Journal of medicinal chemistry*, 54:4880–4895, 2011.

[73] **R. Gurfinkel and T.R. Joy**. Anacetrapib: Hope for CETP inhibitors? *Cardiovascular therapeutics*, 29:327–339, 2010.

[74] **H. Okamoto, F. Yonemori, K. Wakitani, T. Minowa, K. Maeda, and H. Shinkai**. A cholesteryl ester transfer protein inhibitor attenuates atherosclerosis in rabbits. *Nature*, 406:203–207, 2000.

[75] **M. Ranalletta, K.K. Bierilo, Y. Chen, D. Milot, Q. Chen, E. Tung, C. Houde, N.H. Elowe, M. Garcia-Calvo, G. Porter, S. Eveland, B. Frantz-Wattley, M. Kavana, G. Addona, P. Sinclair, C. Sparrow, E.A. O'Neill, K.S. Koblan, A. Sitlani, B. Hubbard, and T.S. Fisher**. Biochemical characterization of cholesteryl ester transfer protein inhibitors. *Journal of lipid research*, 51:2739–2752, 2010.

[76] **S. Kumar, E.Y. Tan, G. Hartmann, Z. Biddle, A.J. Bergman, J. Dru, J.Z. Ho, A.N. Jones, S.J. Staskiewicz, M.P. Braun, B. Karanam, D.C. Dean, I.N. Gendrano, M.W. Graves, J.A. Wagner, and R. Krishna**. Metabolism and excretion of anacetrapib, a novel inhibitor of the cholesteryl ester transfer protein, in humans. *Drug Metabolism and Disposition*, 38:474–483, 2010.

[77] **E.Y. Tan, G. Hartmann, Q. Chen, A. Pereira, S. Bradley, G. Doss, A.S. Zhang, J.Z. Ho, M.P. Braun, D.C. Dean, W. Tang, and S. Kumar**. Pharmacokinetics, metabolism, and excretion of anacetrapib, a novel inhibitor of the cholesteryl ester transfer protein, in rats and rhesus monkeys. *Drug Metabolism and Disposition*, 38:459–473, 2010.

[78] **R. Krishna, A.J. Bergman, B. Jin, A. Garg, B. Roadcap, R. Chiou, J. Dru, J. Cote, T. Laethem, R.W. Wang, V. Didolkar, E. Vets, K. Gottesdiener, and J.A. Wagner**. Assessment of the CYP3A-mediated drug interaction potential of anacetrapib, a potent cholesteryl ester transfer protein (CETP) inhibitor, in healthy volunteers. *The Journal of Clinical Pharmacology*, 49:80–87, 2009.

[79] **R. Krishna, M.S. Anderson, A.J. Bergman, B. Jin, M. Fallon, J. Cote, K. Rosko, C. Chavez-Eng, R. Lutz, D.M. Bloomfield, M. Gutierrez, J. Doherty, F. Bieberdorf, J. Chodakewitz, K.M Gottesdiener, and J.A. Wagner**. Effect of the cholesteryl ester transfer protein inhibitor, anacetrapib, on lipoproteins in patients with dyslipidaemia and on 24-h ambulatory blood pressure in healthy individuals: Two double-blind, randomised placebo-controlled phase I studies. *The Lancet*, 370:1907–1914, 2007.

[80] **D. Bloomfield, G.L. Carlson, A. Sapre, D. Tribble, J.M. McKenney, T.W. Littlejohn III, C.M.C. Sisk, Y. Mitchel, and R.C. Pasternak**. Efficacy and safety of the cholesteryl ester transfer protein inhibitor anacetrapib as monotherapy and coadministered with atorvastatin in dyslipidemic patients. *American heart journal*, 157:352–360, 2009.

[81] **C.P. Cannon, H.M. Dansky, M. Davidson, A.M. Gotto Jr, E.A. Brinton, A.L. Gould, M. Stepanavage, S.X. Liu, S. Shah, J. Rubino, P. Gibbons, A. Hermanowski-Vosatka, B. Binkowitz, Y. Mitchel, and P. Barter**. Design of the DEFINE trial: Determining the EFficacy and tolerability of CETP INhibition with anacEtrapib. *American heart journal*, 158:513–519, 2009.

[82] **L. Yvan-Charvet, F. Matsuura, N. Wang, M.J. Bamberger, T. Nguyen, F. Rinninger, X.C. Jiang, C.L. Shear, and A.R. Tall**. Inhibition of cholesteryl ester transfer protein by torcetrapib modestly increases macrophage cholesterol efflux to HDL. *Arteriosclerosis, thrombosis, and vascular biology*, 27:1132–1138, 2007.

[83] **L. Yvan-Charvet, J. Kling, T. Pagler, H. Li, B. Hubbard, T. Fisher, C.P. Sparrow, A.K. Taggart, and A.R. Tall**. Cholesterol efflux potential and anti-inflammatory properties of high-density lipoprotein after treatment with niacin or anacetrapib. *Arteriosclerosis, thrombosis, and vascular biology*, 30:1430–1438, 2010.

[84] **M.E. Brousseau, E.J. Schaefer, M.L. Wolfe, L.A.T. Bloedon, A.G. Digenio, R.W. Clark, J.P. Mancuso, and D.J. Rader**. Effects of an inhibitor of cholesteryl ester transfer protein on HDL cholesterol. *New England Journal of Medicine*, 350:1505–1515, 2004.

[85] **R.W. Clark, T.A. Sutfin, R.B. Ruggeri, A.T. Willauer, E.D. Sugarman, G. Magnus-Aryitey, P.G. Cosgrove, T.M. Sand, R.T. Wester, J.A. Williams, M.E. Perlman, and M.J. Bamberger**. Raising high-density lipoprotein in humans through inhibition of cholesteryl ester transfer protein. *Arteriosclerosis, thrombosis, and vascular biology*, 24:490–497, 2004.

[86] **S.E. Nissen, J.C. Tardif, S.J. Nicholls, J.H. Revkin, C.L. Shear, W.T. Duggan, W. Ruzyllo, W.B. Bachinsky, G.P. Lasala, and E.M. Tuzcu**. Effect of torcetrapib on the progression of coronary atherosclerosis. *New England Journal of Medicine*, 356:1304–1316, 2007.

[87] **M.L. Bots, F.L. Visseren, G.W. Evans, W.A. Riley, J.H. Revkin, C.H. Tegeler, C.L. Shear, W.T. Duggan, R.M. Vicari, D.E. Grobbee,**

and J.J. Kastelein. Torcetrapib and carotid intima-media thickness in mixed dyslipidaemia (RADIANCE 2 study): A randomised, double-blind trial. *The Lancet*, 370:153–160, 2007.

[88] **J.J.P. Kastelein, S.I. van Leuven, L. Burgess, G.W. Evans, J.A. Kuivenhoven, P.J. Barter, J.H. Revkin, D.E. Grobbee, W.A. Riley, C.L. Shear, W.T. Duggan, and M.L. Bots**. Effect of torcetrapib on carotid atherosclerosis in familial hypercholesterolemia. *New England Journal of Medicine*, 356:1620–1630, 2007.

[89] **S.J. Nicholls, E.M. Tuzcu, D.M. Brennan, J.C. Tardif, and S.E. Nissen**. Cholesteryl ester transfer protein inhibition, high-density lipoprotein raising, and progression of coronary atherosclerosis: Insights from ILLUSTRATE (investigation of lipid level management using coronary ultrasound to assess reduction of atherosclerosis by CETP inhibition and HDL elevation). *Circulation*, 118:2506–2514, 2008.

[90] **L. Xie, J. Li, L. Xie, and P.E. Bourne**. Drug discovery using chemical systems biology: Identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS computational biology*, 5:e1000387, 2009.

[91] **H. Shinkai, K. Maeda, T. Yamasaki, H. Okamoto, and I. Uchida**. bis (2-(acylamino) phenyl) disulfides, 2-(acylamino) benzenethiols, and s-(2-(acylamino) phenyl) alkanethioates as novel inhibitors of cholesteryl ester transfer protein. *Journal of medicinal chemistry*, 43:3566–3572, 2000.

[92] **M. Vergeer and E.S.G. Stroes**. The pharmacology and off-target effects of some cholesterol ester transfer protein inhibitors. *The American journal of cardiology*, 104:32E–38E, 2009.

[93] **G.J. de Grooth, J.A. Kuivenhoven, A.F.H. Stalenhoef, J. de Graaf, A.H. Zwinderman, J.L. Posma, A. van Tol, and J.J.P. Kastelein**. Efficacy and safety of a novel cholesteryl ester transfer protein inhibitor, JTT-705, in humans: A randomized phase II dose-response study. *Circulation*, 105: 2159–2165, 2002.

[94] **J.A. Kuivenhoven, G.J. de Grooth, H. Kawamura, A.H. Klerkx, F. Wilhelm, M.D. Trip, and J.J.P. Kastelein**. Effectiveness of inhibition of cholesteryl ester transfer protein by JTT-705 in combination with pravastatin in type II dyslipidemia. *The American journal of cardiology*, 95:1085–1088, 2005.

[95] **E.A. Stein, E.S.G. Stroes, G. Steiner, B.M. Buckley, A.M. Capponi, T. Burgess, E.J. Niesor, D. Kallend, and J.J.P. Kastelein**. Safety and tolerability of dalcetrapib. *The American journal of cardiology*, 104:82–91, 2009.

[96] **Hoffmann-La Roche**. A study of ro4607381 in stable coronary heart disease patients with recent acute coronary syndrome. `http://clinicaltrials.gov/ct2/show/NCT00658515`. Retrieved on September 5, 2011.

[97] **B. Hess, D. van der Spoel, E. Lindahl, E. Apol, R. Apostolov, H.J.C. Berendsen, A. van Buuren, P. Bjelkmar, R. van Drunen, A. Feenstra, G. Groenhof, P. Kasson, P. Larsson, P. Meulenhoff, T. Murtola, S. Pall, S. Pronk, R. Schulz, M. Shirts, A. Sijbers, and P. Tieleman**. Gromacs user manual version 4.5.4. `http://www.gromacs.org`, 2010. Retrieved on September 19, 2011.

[98] **T. Schlick**. *Molecular modeling and simulation: An interdisciplinary guide*. Springer, New York, 2006.

[99] **D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J.C. Berendsen**. Gromacs: Fast, flexible, and free. *Journal of computational chemistry*, 26:1701–1718, 2005.

[100] **R.W. Hockney, S.P. Goel, and J.W. Eastwood**. Quiet high-resolution computer models of a plasma. *Journal of Computational Physics*, 14:148–158, 1974.

[101] **H.J.C. Berendsen and W.F. van Gunsteren**. Practical algorithms for dynamics simulations. *Molecular simulation*, 1:173–185, 1988.

[102] **M.P. Allen and D.J. Tildesley**. *Computer simulation of liquids*. Clarendon Press, 1999.

[103] **H.J.C. Berendsen, J.P.M. Postma, W.F. Van Gunsteren, A. DiNola, and J.R. Haak**. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81:3684–3690, 1984.

[104] **W.G. Hoover**. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31:1695–1697, 1985.

[105] **M. Parrinello and A. Rahman**. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52:7182–7190, 1981.

[106] **A.Y. Toukmaji and J.A. Board**. Ewald summation techniques in perspective: A survey. *Computer Physics Communications*, 95:73–92, 1996.

[107] **P.P. Ewald**. Ewald summation. *Ann. Phys*, 369, 1921.

[108] **B. Hess, H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije**. LINCS: A linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18:1463–1472, 1997.

[109] **V. Kräutler, W.F. van Gunsteren, and P.H. Hünenberger**. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*, 22: 501–508, 2001.

[110] **H.C. Andersen**. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52:24–34, 1983.

[111] **W.L. Jorgensen, D.S. Maxwell, and J. Tirado-Rives**. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118: 11225–11236, 1996.

[112] **O. Trott and A.J. Olson**. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31:455–461, 2010.

[113] **B.L. Dong, Q.H. Liao, and J. Wei**. Docking and molecular dynamics study on the inhibitory activity of n, n-disubstituted-trifluoro-3-amino-2-propanols-based inhibitors of cholesteryl ester transfer protein. *Journal of Molecular Modeling*, 17:1727–1734, 2010.

[114] **W. Humphrey, A. Dalke, and K. Schulten**. Vmd: Visual molecular dynamics. *Journal of molecular graphics*, 14:33–38, 1996.

[115] **G. Vrient**. DSSP. `http://swift.cmbi.ru.nl/gv/dssp/`. Retrieved on October 7, 2011.

[116] **Rizzo Research Lab**. MD simulation: Protein in water. `http://ringo.ams.sunysb.edu/index.php/MD_Simulation:_Protein_in_Water`. Retrieved on October 7, 2011.

[117] **W. Kabsch and C. Sander**. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637, 1983.

[118] **V.N. Maiorov and G.M. Crippen**. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *Journal of Molecular Biology*, 235:625–634, 1994.

[119] **B. de Groot**. Principal component analysis. `http://www.mpibpc.mpg.de/groups/de_groot/compbio1/p4/index.html#intro`. Retrieved on October 6, 2011.

[120] **A. Chatterjee and A.K. Bothra**. Molecular dynamics simulation of Angiotensin I: Precursor of Angiotensin II. *Int. J. Integ. Biol*, 9:21–25, 2010.

[121] **Z.S. Derewenda, L. Lee, and U. Derewenda**. The occurence of CH-O hydrogen bonds in proteins. *Journal of molecular biology*, 252:248–262, 1995.

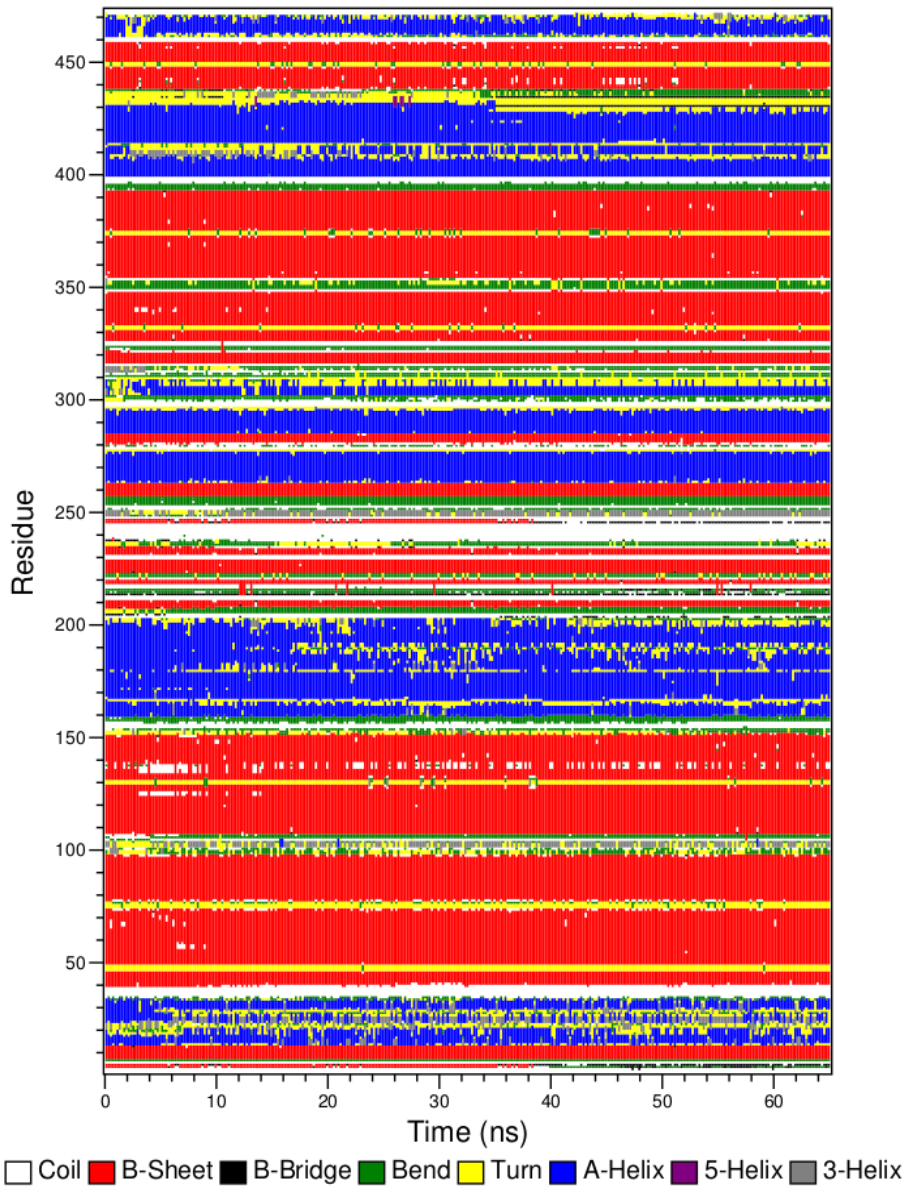
[122] **P.B. Balbuena and J.M. Seminario**. *Molecular dynamics: from classical to quantum methods*, volume 7. Elsevier Science, 1999.

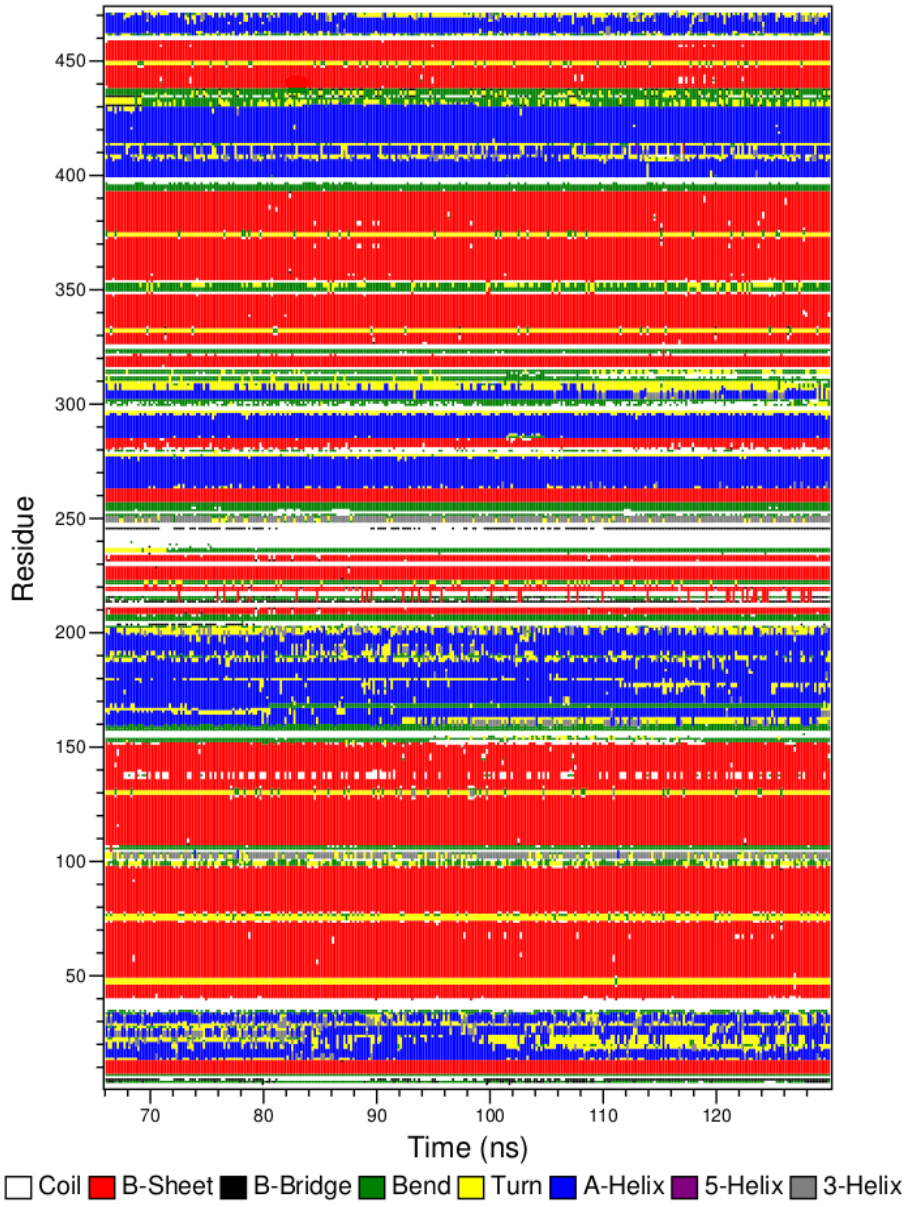
[123] **T.V. Bogdan and E.V. Isaeva**. Molecular dynamic simulation of the structure of liquid chlorobenzene, *ortho*-chlorotoluene and their mixtures. *Journal of Structural Chemistry*, 50:640–646, 2009.

A. APPENDIX

A.1 Secondary structures of systems simulated for 200 ns

A.1.1 L1-200ns




Residue
450
400
350
300
250
200
150
100
50
70
80
90
100
110
120
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix

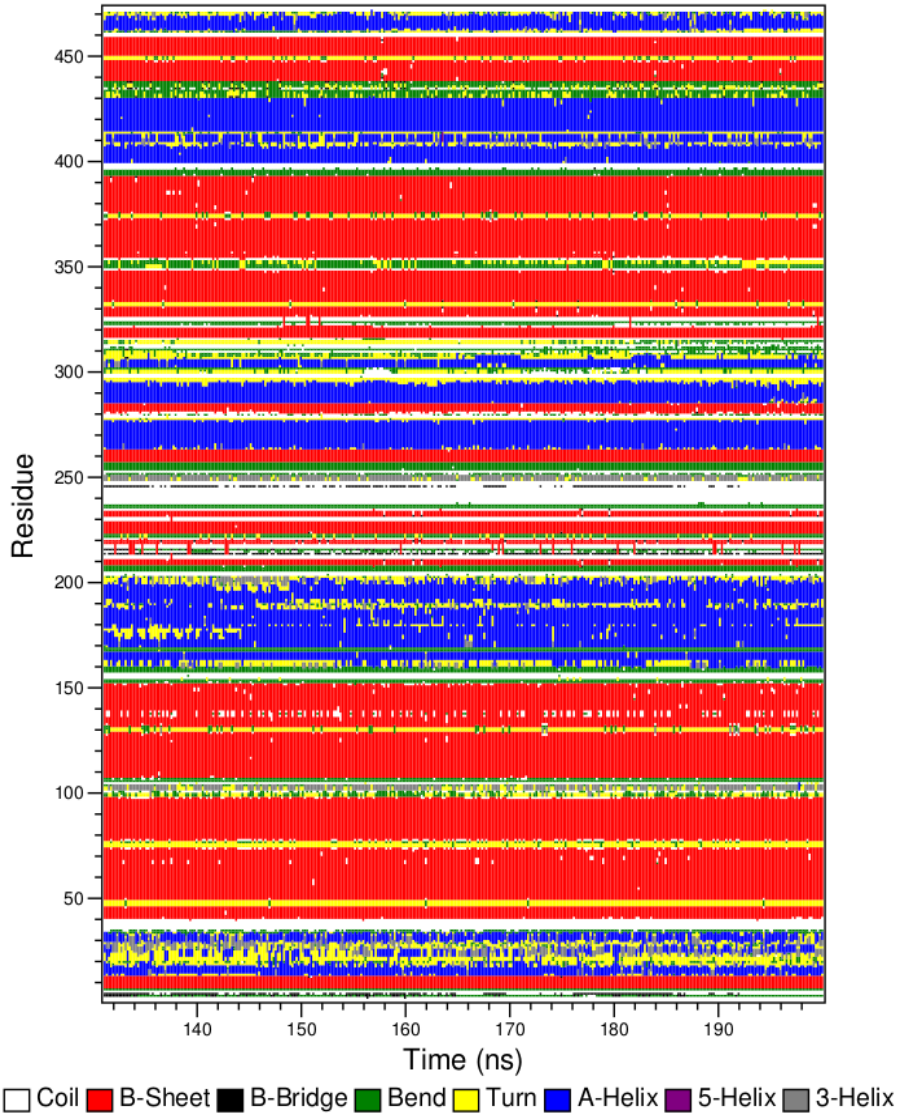
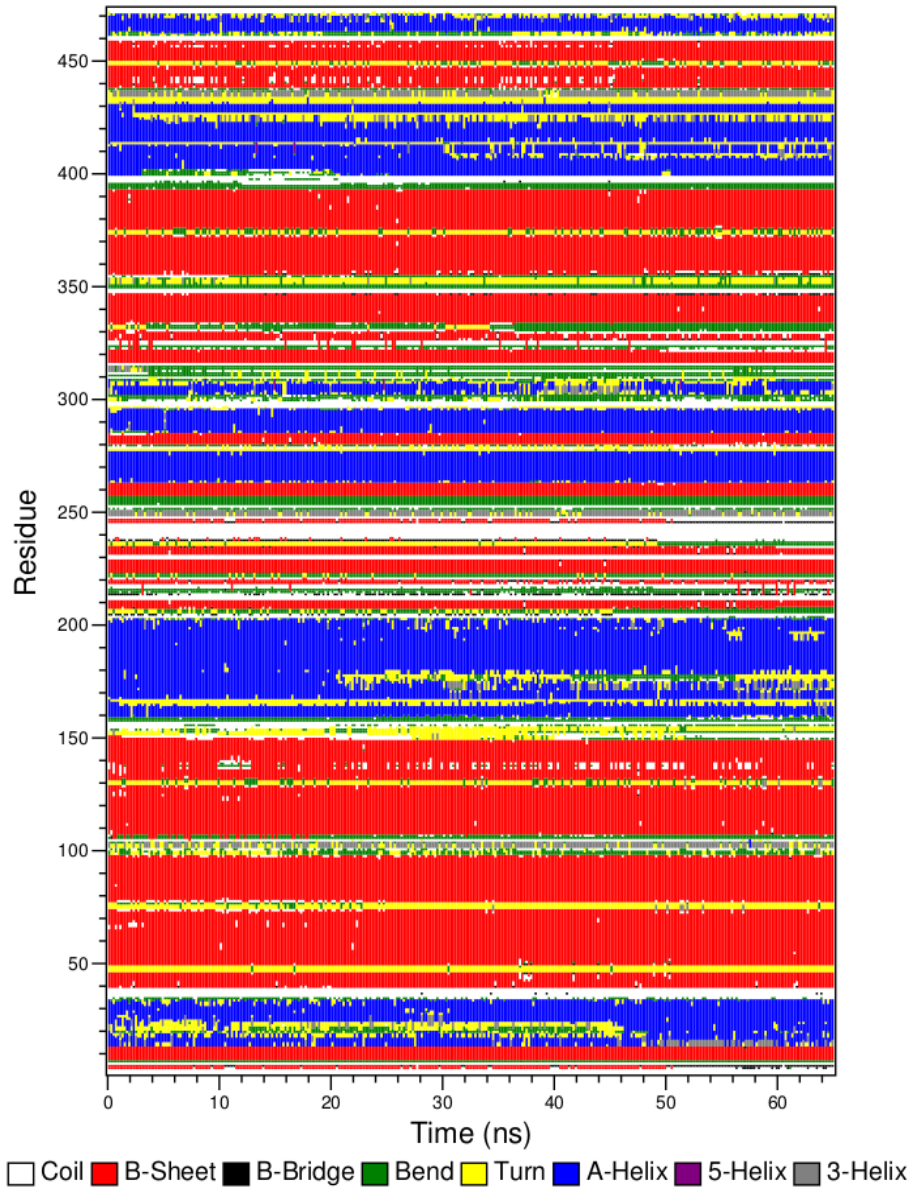
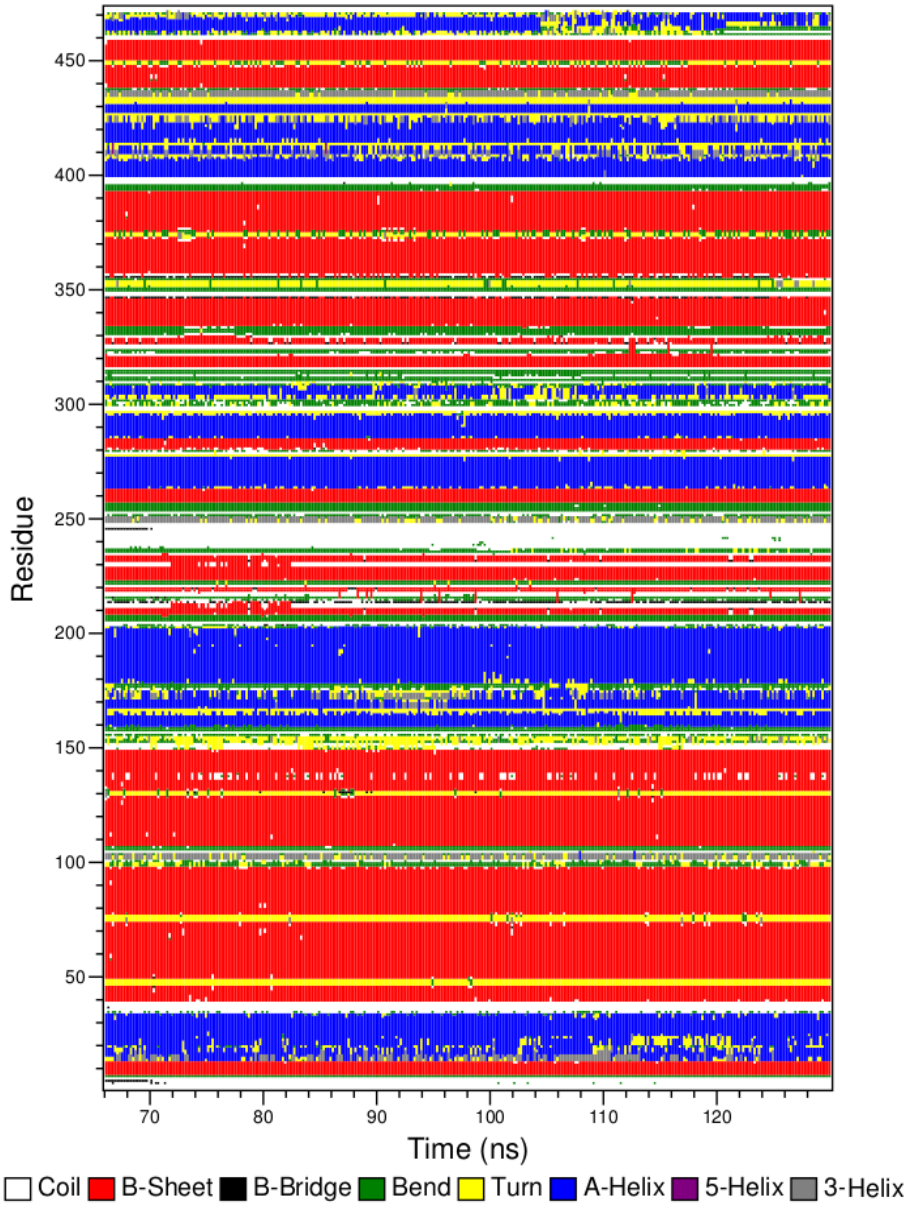


Figure A.1: Secondary structure of CETP during the simulation L1-200ns. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.1.2 L2-200ns


Residue
450
400
350
300
250
200
150
100
50
0
10
20
30
40
50
60
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix


Residue
450
400
350
300
250
200
150
100
50
70
80
90
100
110
120
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix

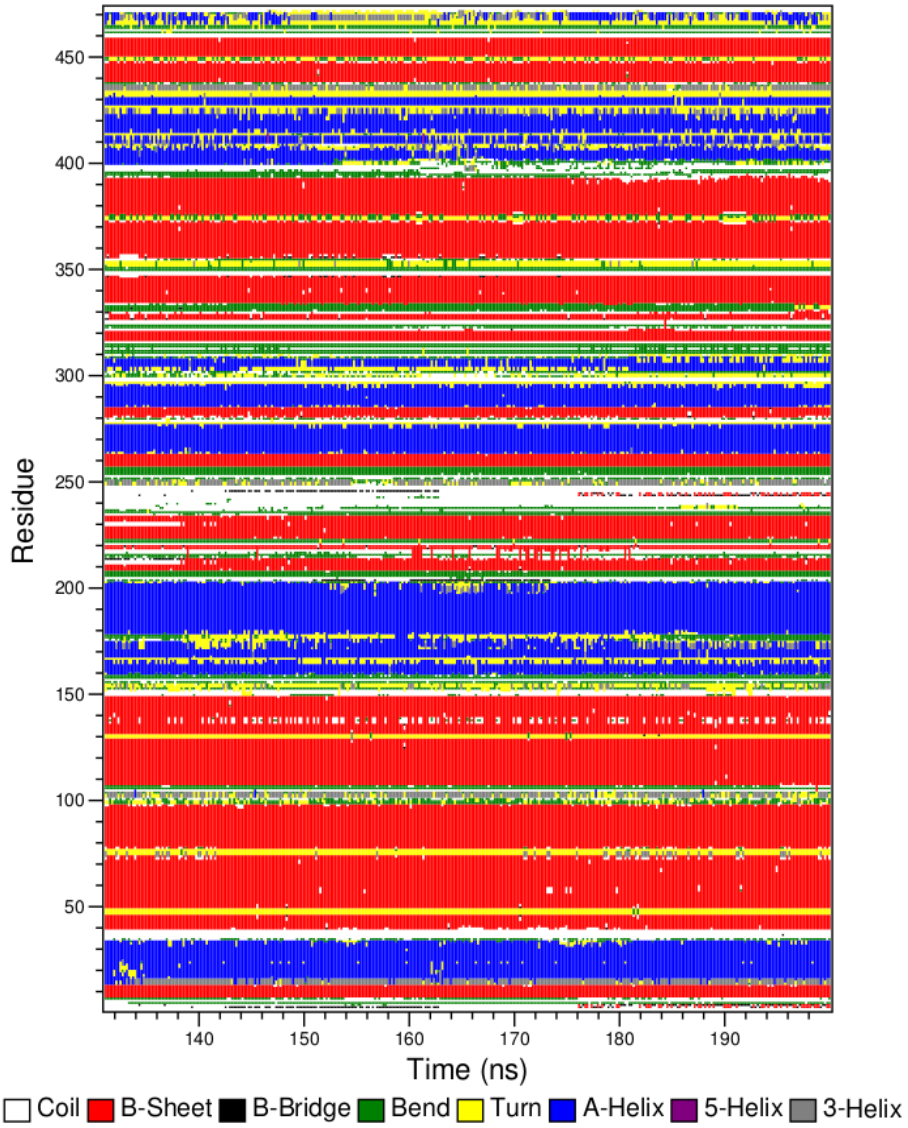
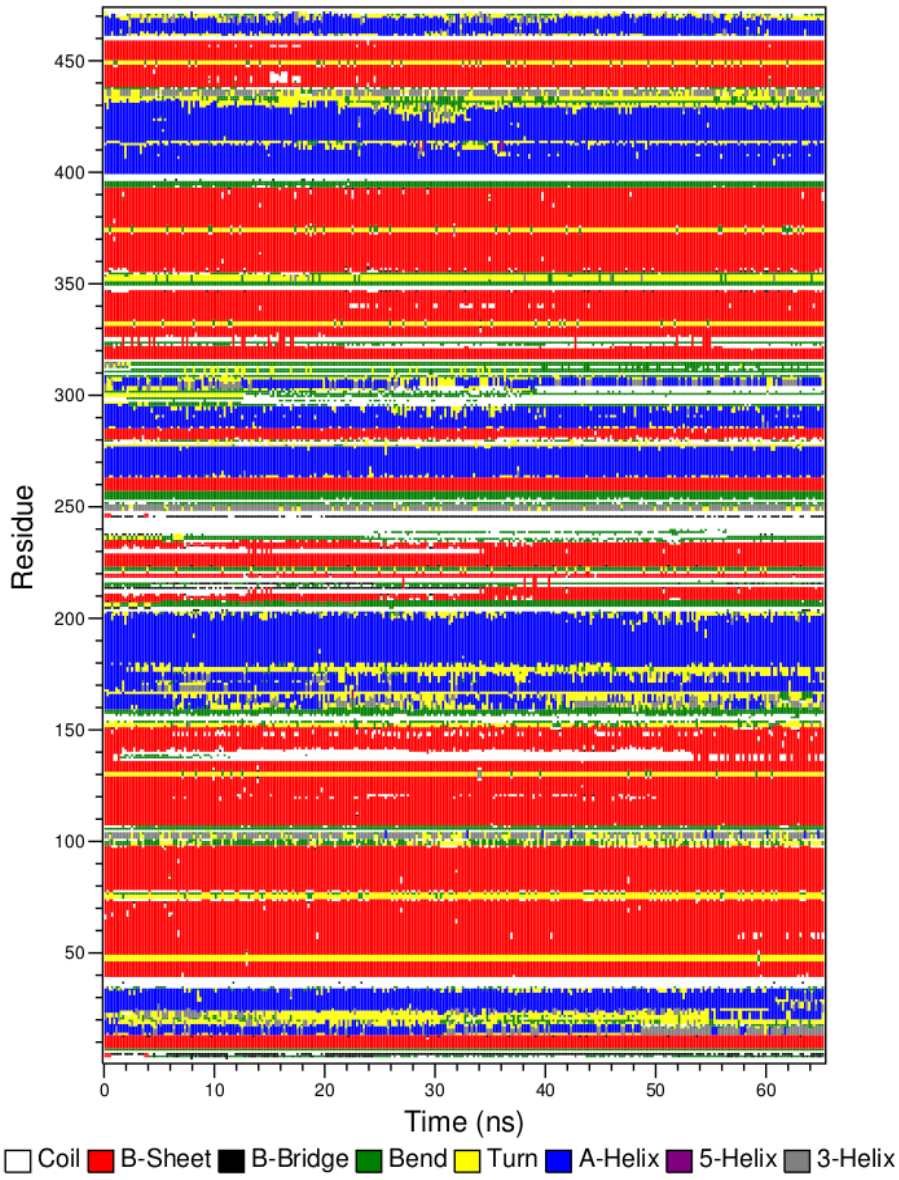
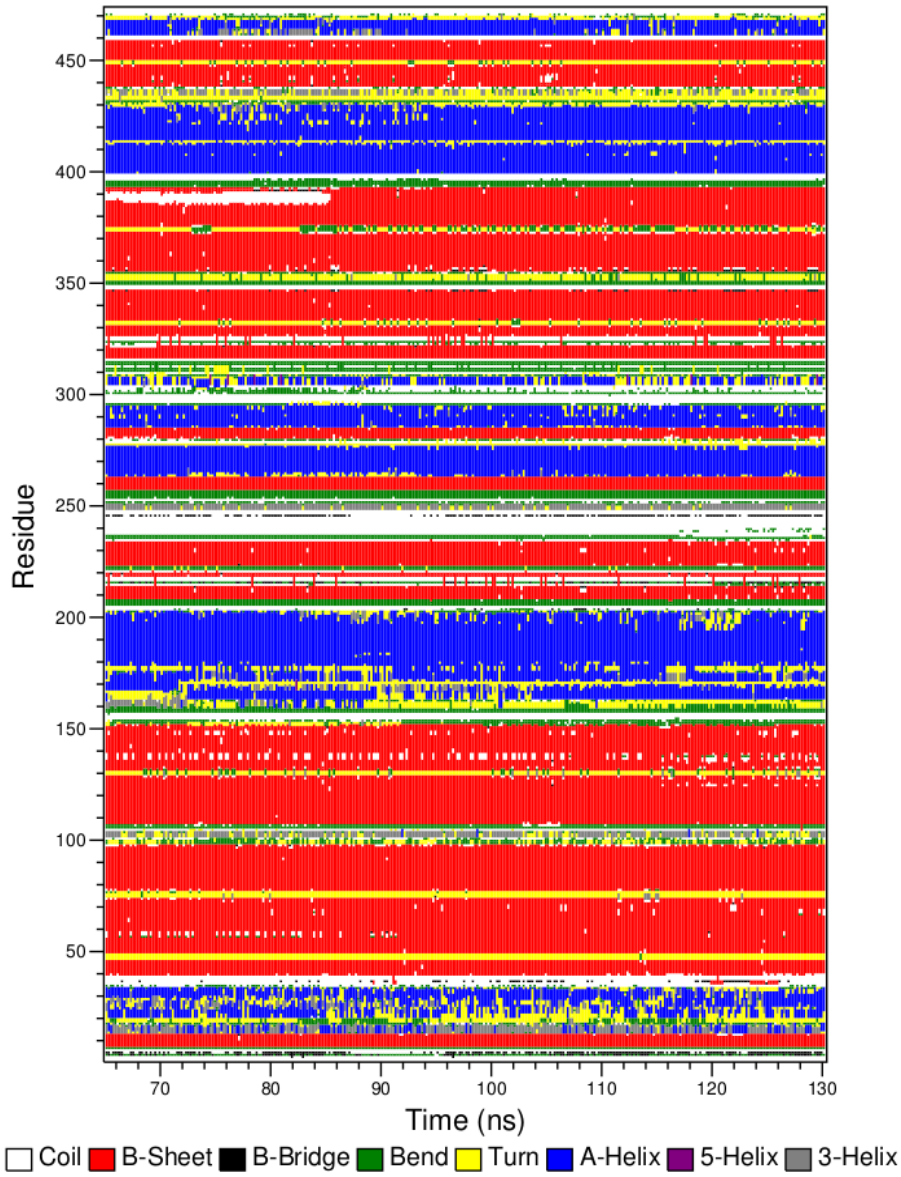


Figure A.2: Secondary structure of CETP during the simulation L2-200ns. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.1.3 L3-200ns


Residue
450
400
350
300
250
200
150
100
50
0
10
20
30
40
50
60
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix


Residue
450
400
350
300
250
200
150
100
50
70
80
90
100
110
120
130
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix

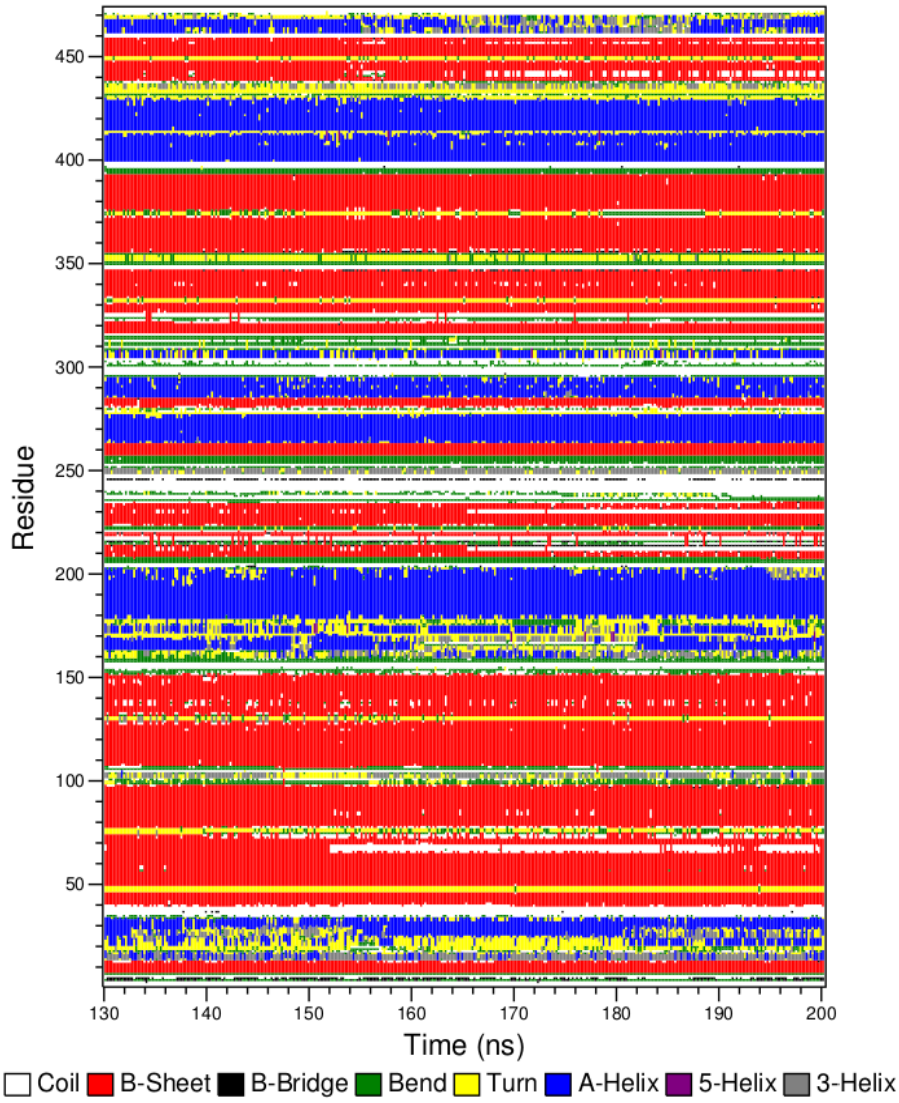
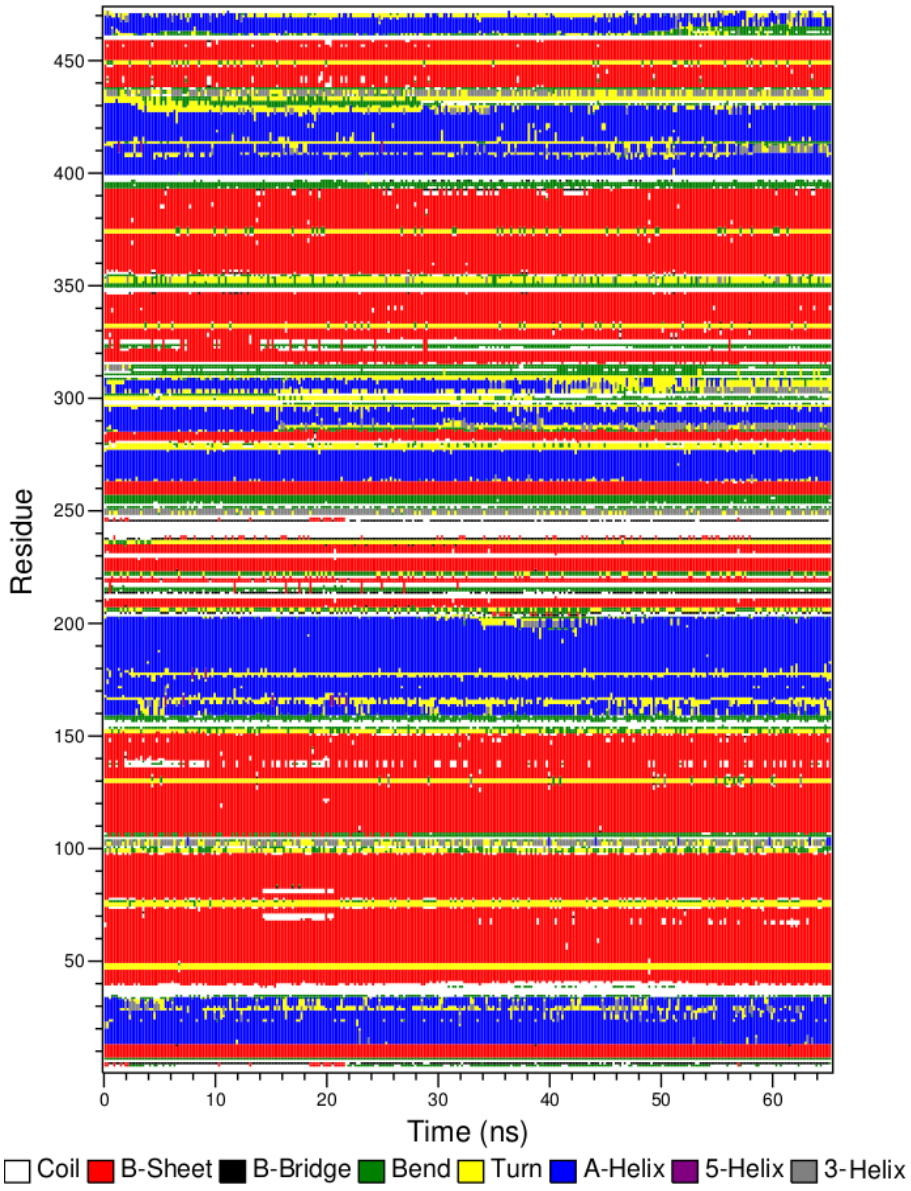
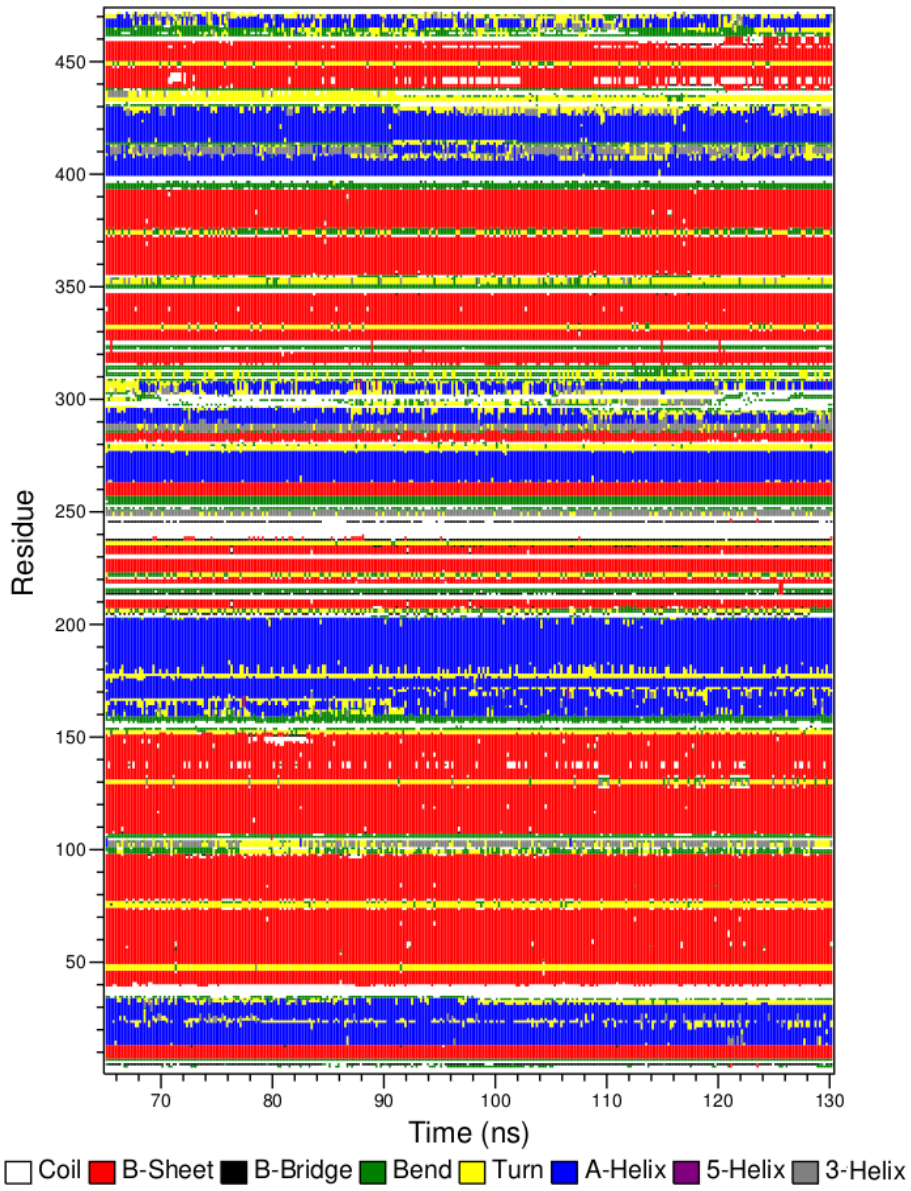


Figure A.3: Secondary structure of CETP during the simulation L3-200ns. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.1.4 L4-200ns


Residue
450
400
350
300
250
200
150
100
50
0
10
20
30
40
50
60
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix


Residue
450
400
350
300
250
200
150
100
50
70
80
90
100
110
120
130
Time (ns)
Coil
B-Sheet
B-Bridge
Bend
Turn
A-Helix
5-Helix
3-Helix

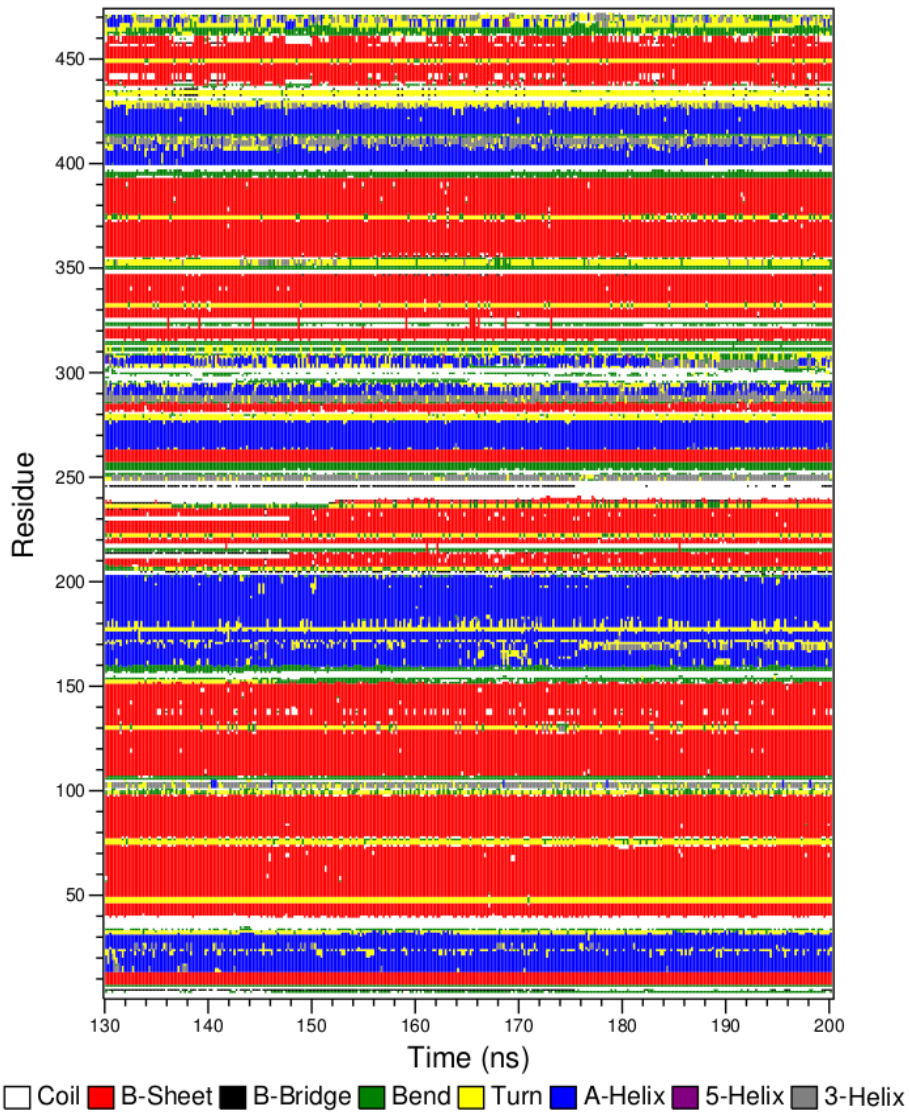


Figure A.4: Secondary structure of CETP during the simulation L4-200ns. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2 Secondary structures of systems simulated for 20 ns

A.2.1 S1-helix

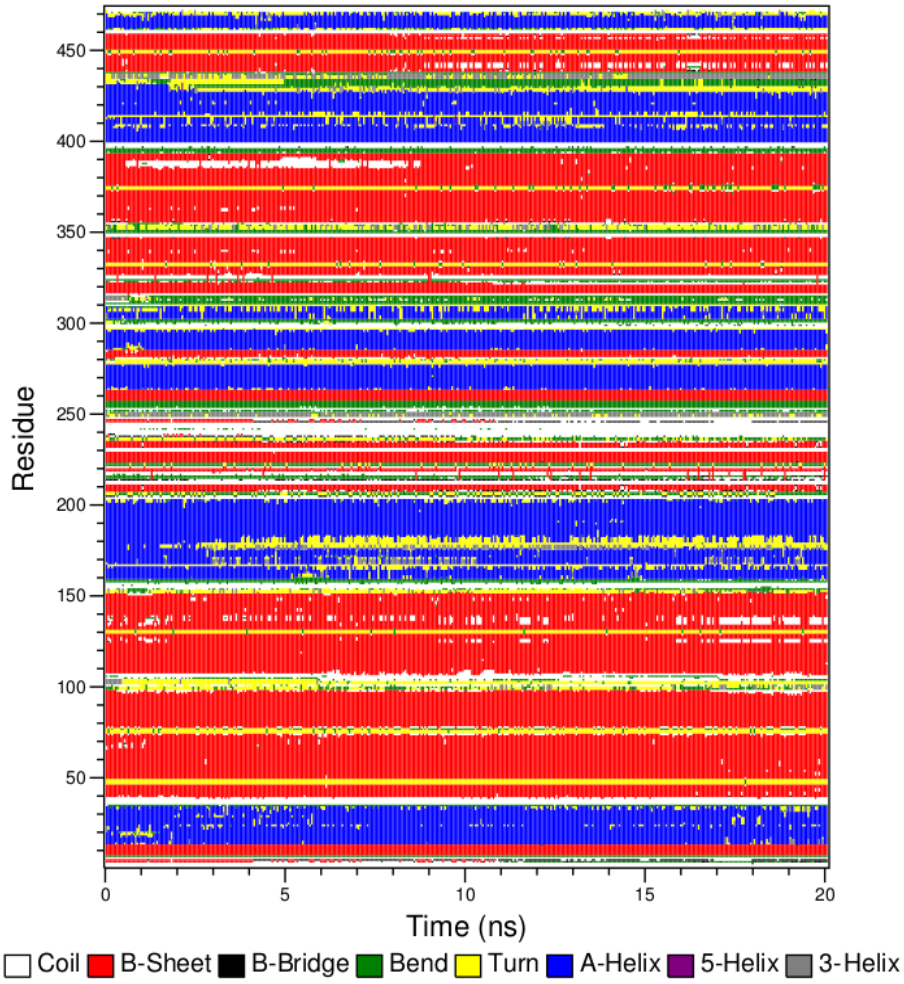


Figure A.5: Secondary structure of CETP during the simulation S1-helix. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2.2 S2-1nm

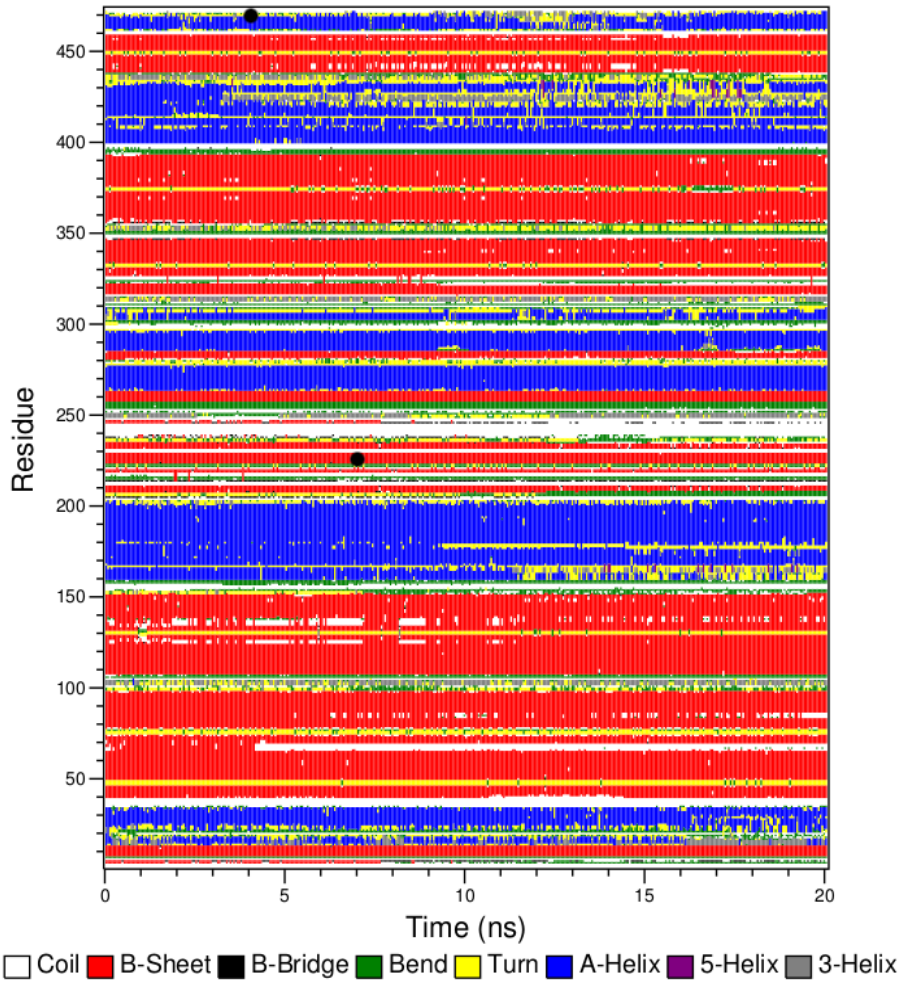


Figure A.6: Secondary structure of CETP during the simulation S2-1nm. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2.3 S3-2nm

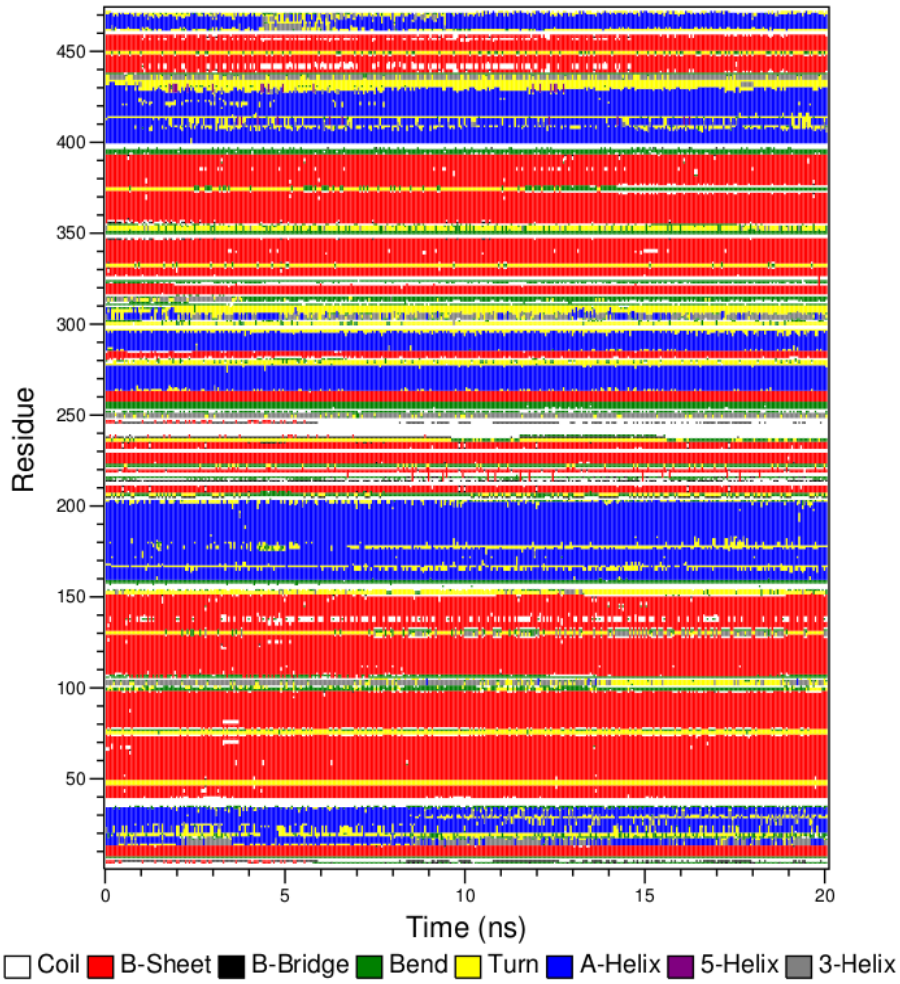


Figure A.7: Secondary structure of CETP during the simulation S3-2nm. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2.4 S4-3nm

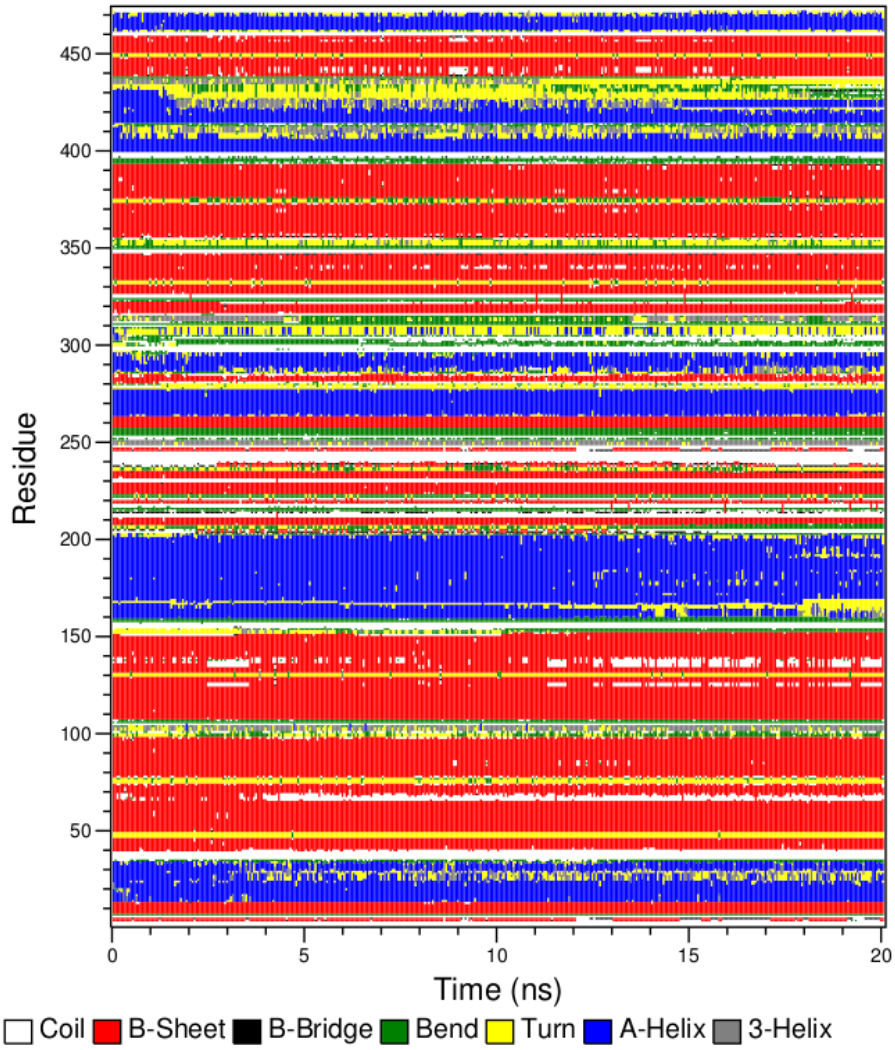


Figure A.8: Secondary structure of CETP during the simulation S4-3nm. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2.5 S5-4nm

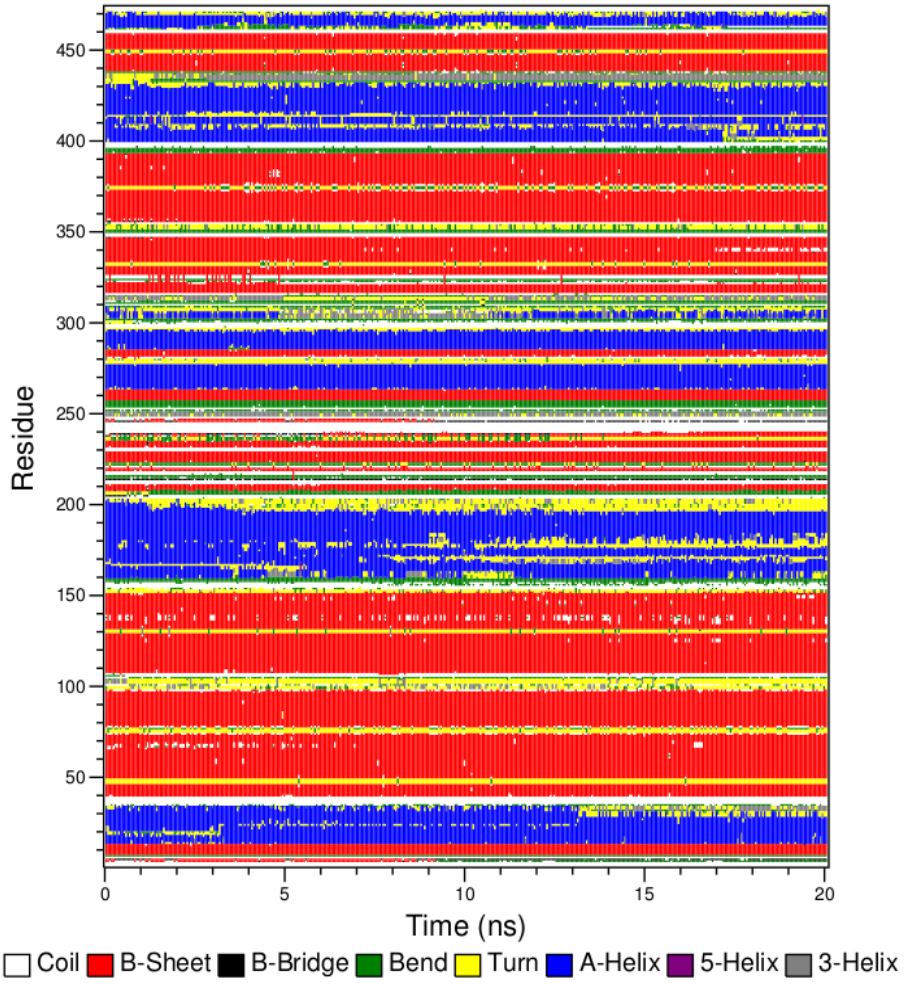


Figure A.9: Secondary structure of CETP during the simulation S5-4nm. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.

A.2.6 S6-convex

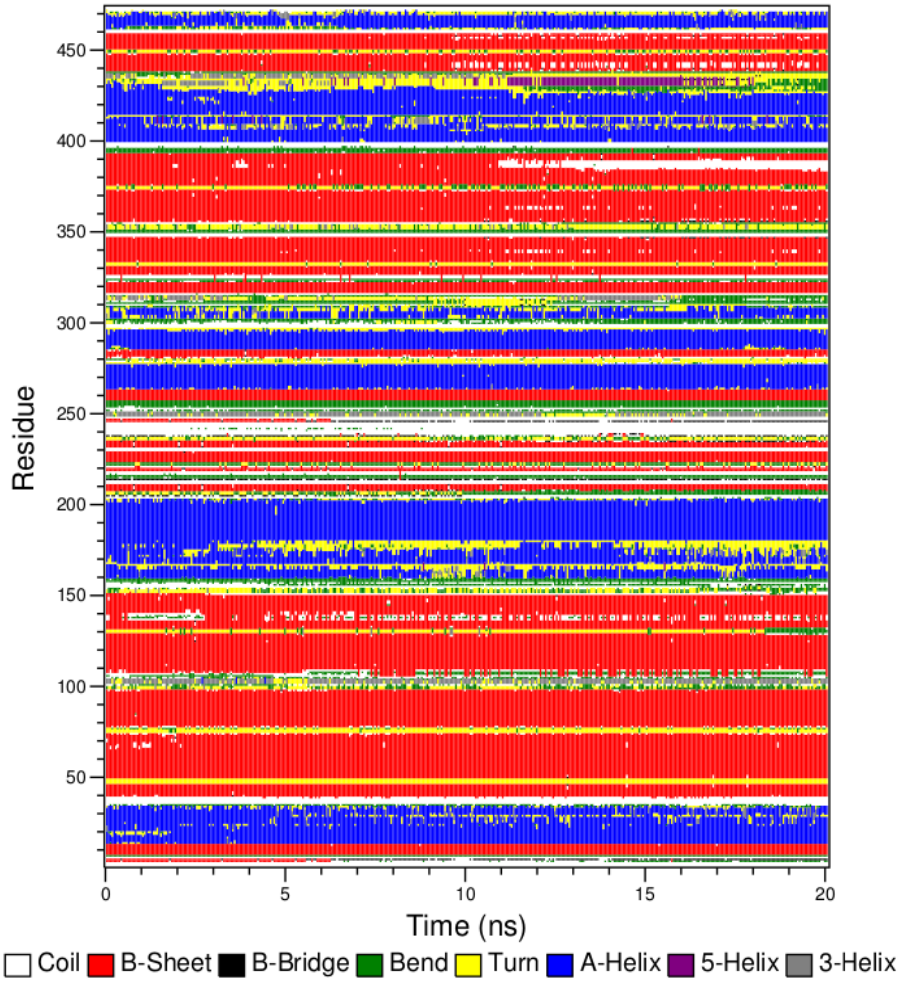


Figure A.10: Secondary structure of CETP during the simulation S6-convex. The secondary structure does not change greatly during the simulation despite the small fluctuations. The changes within helix X (residues 461-472) are of great importance.